

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:44:21 ; Search time 450.327 Seconds  
(without alignments)  
15572.506 Million cell updates/sec

Title: US-09-766-511b-51

Perfect score: 3114  
Sequence: 1 cttatgttggaagtcctt.....tttaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:
	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3114	100.0	3114	22	AAS01375 Human TANGO 405 cD
2	591.4	19.0	850	20	AAZ07531 Human SDCMP3 poly
3	379.2	12.2	1227	19	AAV42549 Mouse dectin-2 cDN
4	379.2	12.2	1252	22	AAS01378 Murine TANGO 405 a
5	376.2	12.1	630	20	AAZ07532 Mouse SDCMP3 poly
6	364.6	11.7	3708	20	AAZ77530 Human ovarian tumo
7	364.6	11.7	3748	21	AAZ77686 Human cancer assoc
8	356.2	11.4	3647	24	ABK83658 Human cDNA differe
9	351.4	11.3	803	22	AAI97506 Human neuroblastom

c 11	345.8	11.1	821	22	AAS01376 Murine TANGO 405 c
c 11	317.6	10.2	56583	21	AAF21125 Human low adenosin
c 12	317.6	10.2	56583	21	AA335003 Human adenosine re
c 13	303.2	9.7	501	19	AAV42554 Mouse dectin-2 ext
14	283.8	9.1	1312	22	AAF90241 Nucleotide sequenc
15	272.8	8.8	827	22	AAI9729 Dendritic cell (DC
16	269	8.6	800	22	AAI9730 Dendritic cell (DC
c 17	257.8	8.3	3298	17	AAI45082 Beta-1-4-galactosy
c 18	255.4	8.2	684	22	AAF98678 Human ovarian canc
c 19	254.2	8.2	324	22	AAK63162 Human immune/haema
c 20	244.4	7.2	3929	24	AB190324 Mouse ischaemic co
21	201.2	6.5	444	24	ABK52914 Human cDNA encodin
22	200.4	6.4	402	24	ABK52903 Human cDNA encodin
23	197	6.3	1091	22	AA331385 Human cDNA encodin
24	197	6.3	1091	24	ABQ66709 Human polynucleoti
25	197	6.3	1096	22	AA331224 Human cDNA encodin
26	197	6.3	1096	24	ABQ66548 Human polynucleoti
27	197	6.3	1104	20	AA04865 Primate DCMPI C-le
c 28	197	6.3	1307	24	AB190686 Human polynucleoti
c 29	179.4	5.8	582	24	ABK52904 Human dendritic ce
c 30	177.8	5.7	753	24	ABK52906 Human dendritic ce
c 31	169.6	5.4	1013	24	ABK52905 Human dendritic ce
c 32	154.8	5.0	558	24	ABK52907 Human dendritic ce
33	144.8	4.6	561	24	ABK52911 Human dendritic ce
34	140	4.5	448	24	ABK52912 Human dendritic ce
c 35	140	4.5	666	24	ABK52910 Human dendritic ce
c 36	137.4	4.4	568	24	ABK52908 Human dendritic ce
c 37	137.4	4.4	1418	20	AA04867 rodent DCMPI C-lec
c 38	130.2	4.2	403	24	ABK52909 Human dendritic ce
c 39	118.6	3.8	375	16	AAI24623 Human gene signatu
c 40	117.4	3.8	968	22	AAE81743 Human membrane ass
41	117.4	3.8	997	20	AA52274 Protein PRO244 cDN
42	117.4	3.8	997	22	AA545936 Human DNA encoding
43	117.4	3.8	997	22	AAI72432 Human PRO244 cDNA.
c 44	111.8	3.6	458	22	AA524836 Human ovarian PCR-
c 45	111.8	3.6	567	22	AAH83473 Human ovarian tumo

#### ALIGNMENTS

#### RESULT 1

AAS01375  
ID AAS01375 standard; cDNA; 3114 BP.

XX AAS01375;

XX AC AAS01375;

XX DT 04-JUL-2001 (first entry)

XX DE Human TANGO 405 cDNA sequence.

XX KW Human TANGO 210; clone jthLa152h06; TANGO 364; TANGO 366; dectin-2;  
INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
gene therapy; growth modulator; proliferation; cell differentiation;  
lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
autoimmune disorder; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX CDS 154..783

XX FT /\*tag= a

XX FT /product= "TANGO 405 protein"

XX FT /note= "The ORF is specifically claimed"

XX FT 154..297

XX FT /\*tag= b

XX FT 298..780

XX FT /\*tag= c

XX PN WO200118016-A1.

XX PD 15-MAR-2001.

PF 30-JUN-2000; 2000WO-US18174.  
 XX  
 PR 10-SEP-1999; 99US-0393996.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
 XX  
 DR WPI; 2001-183280/18.  
 XX P-PSDB; AAU00479.  
 XX  
 PT Isolated nucleic acid molecules encoding proteins useful as modulating  
 PT agents in regulating a variety of cellular processes are used for  
 PT treating e.g. cancer and autoimmune disorders -  
 XX  
 PS Claim 2; Fig 6A-6C; 326pp; English.  
 XX  
 CC The present sequence encoding for human TANGO 405 protein is isolated  
 CC from cDNA clone jthla152h06 from a human mixed lymphocyte reaction cDNA  
 CC library. It is 1 of 6 novel human proteins which include TANGO 210  
 CC (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394  
 CC (AAU00473) and INTERCEPT 400 (AAU00476). Novel sequences for murine  
 CC TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and  
 CC a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic  
 CC acids encoding these novel proteins are useful as modulating agents in  
 CC regulating a variety of cellular processes and can be used to express  
 CC the proteins in a host cell in gene therapy applications. Human and  
 CC murine TANGO 405 proteins show sequence homology to murine lectin-2.  
 CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
 CC activity, morphology and movement/migration of human lymphocytes and  
 CC bone marrow cells and tissues and can be used to prevent, diagnose or  
 CC treat leukaemia, lymphomas and autoimmune disorders.  
 XX  
 SO Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 other;  
 XX

Query Match 100.0%; Score 3114; DB 22; Length 3114;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAATGTTGGAAGTCCTTAGTCCTATGAGAGTGTAGCAGTTGTGCCCTGAGCTCTA 60  
 DB 1 CTTAATGTTGGAAGTCCTTAGTCCTATGAGAGTGTAGCAGTTGTGCCCTGAGCTCTA 60

QY 61 GCTTCTTTAAATGAAGCTGAGTCTCTGGCAACATCTTTAGGAGAGAGGTCACAAAGGT 120  
 DB 61 GCTTCTTTAAATGAAGCTGAGTCTCTGGCAACATCTTTAGGAGAGAGGTCACAAAGGT 120

QY 121 TCCTGGACCTTCTCAACACAGGGAGGCTGCATATGATGCAAGACAGCAACCTCAAAGT 180  
 DB 121 TCCTGGACCTTCTCAACACAGGGAGGCTGCATATGATGCAAGACAGCAACCTCAAAGT 180

QY 181 ACAGAAAAGAGCTGGTGTCCCTGAGACTCTGCTGTGGTGGGATTTCCATTGCA 240  
 DB 181 ACAGAAAAGAGCTGGTGTCCCTGAGACTCTGCTGTGGTGGGATTTCCATTGCA 240

QY 241 CTCCTCAGTGTCTTCATGTGAGCTGTGAGTAACCTTACCATTTACATATGGTCAA 300  
 DB 241 CTCCTCAGTGTCTTCATGTGAGCTGTGAGTAACCTTACCATTTACATATGGTCAA 300

QY 301 ACTGGCAAAGGCTGTCTGAACCTACACTCATATCATTCAGTCTCACTGCTGAGTAA 360  
 DB 301 ACTGGCAAAGGCTGTCTGAACCTACACTCATATCATTCAGTCTCACTGCTGAGTAA 360

QY 361 GGGACAAAGGTGCCAGGCTGGGATGTGCCAGCTCTTGGGAAGTCATTGTGGTCCAGT 420  
 DB 361 GGGACAAAGGTGCCAGGCTGGGATGTGCCAGCTCTTGGGAAGTCATTGTGGTCCAGT 420

QY 421 TGTCTACTTCATTCCAGTGAAGAGAGGTTTGTCTTAAGAGTGAAGCAACTGTGTGAG 480  
 DB 421 TGTCTACTTCATTCCAGTGAAGAGAGGTTTGTCTTAAGAGTGAAGCAACTGTGTGAG 480

QY 481 ATGGAGCACATTGGTGTGTTCAACACAGAGAGCAGCAATTCATTGTCCAGCAG 540  
 DB 481 ATGGAGCACATTGGTGTGTTCAACACAGAGAGCAGCAATTCATTGTCCAGCAG 540

DB 481 ATGGAGCACATTGGTGTGTTCAACACAGAGAGCAGCAATTCATTGTCCAGCAG 540  
 QY 541 CTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCACCAAGGTAATAATAATTGG 600  
 DB 541 CTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCACCAAGGTAATAATAATTGG 600  
 QY 601 CAATGATGATAGACACACCTTATGAGAAAATGTCAAGATTTGGACACCTAGGTAGCC 660  
 DB 601 CAATGATGATAGACACACCTTATGAGAAAATGTCAAGATTTGGACACCTAGGTAGCC 660  
 QY 661 AATCATCTCGACAGCAATGTCTTCAATAGTCCTTGGAAACCTACAGGATGGGCTG 720  
 DB 661 AATCATCTCGACAGCAATGTCTTCAATAGTCCTTGGAAACCTACAGGATGGGCTG 720  
 QY 721 AATGATGTTATCTGTGAACTAGAGGAATTCATATGTGAGATGAATAGATTTACCTA 780  
 DB 721 AATGATGTTATCTGTGAACTAGAGGAATTCATATGTGAGATGAATAGATTTACCTA 780  
 QY 781 TCAGTAGAGCTTAATTCGAAAGAGAGAGCAATTTACTGACGTAATTTTCCCTGACCT 840  
 DB 781 TCAGTAGAGCTTAATTCGAAAGAGAGAGCAATTTACTGACGTAATTTTCCCTGACCT 840  
 QY 841 CTTTAAATTTGAACCTATCATGAAATGATAATTTCTTCTGAAATTTACACATAATCCCT 900  
 DB 841 CTTTAAATTTGAACCTATCATGAAATGATAATTTCTTCTGAAATTTACACATAATCCCT 900  
 QY 901 ATGTTATAGAGTTCACAGAAATGGAAGATACCTGTTCCTTAAATCAATCTTCTCGT 960  
 DB 901 ATGTTATAGAGTTCACAGAAATGGAAGATACCTGTTCCTTAAATCAATCTTCTCGT 960  
 QY 961 TTCTCTTTTCCATTAATGATAGATGACCCCTTCCCTCTCTTTGTCCCATTTCTTCAC 1020  
 DB 961 TTCTCTTTTCCATTAATGATAGATGACCCCTTCCCTCTCTTTGTCCCATTTCTTCAC 1020  
 QY 1021 GTATTCATTTTCTTCTTCCACTTACATACACAAATATTTATTTGTCAGAGACT 1080  
 DB 1021 GTATTCATTTTCTTCTTCCACTTACATACACAAATATTTATTTGTCAGAGACT 1080  
 QY 1081 GTACTATTTTGTGTTAGAAGATTTATAAGGAGATATCTTTTGAAGATTTATGACATTC 1140  
 DB 1081 GTACTATTTTGTGTTAGAAGATTTATAAGGAGATATCTTTTGAAGATTTATGACATTC 1140  
 QY 1141 TTCTCTCAATATACCAATAAGAAATCTTTTGGTCAAGATGGTGGAGACTACAATCAT 1200  
 DB 1141 TTCTCTCAATATACCAATAAGAAATCTTTTGGTCAAGATGGTGGAGACTACAATCAT 1200  
 QY 1201 CTGAAGGCTGACAAAGAGTGTGAAAGACATGTCTTAGATGGCTCACTCACATGGCTG 1260  
 DB 1201 CTGAAGGCTGACAAAGAGTGTGAAAGACATGTCTTAGATGGCTCACTCACATGGCTG 1260  
 QY 1261 AACTTGGTGTGGCTATTAATGTAACCTGGAATAAATTTTATTCGAGTTAGGATTT 1320  
 DB 1261 AACTTGGTGTGGCTATTAATGTAACCTGGAATAAATTTTATTCGAGTTAGGATTT 1320  
 QY 1321 GGCATTTTATATATGTTGATTCATCAAGTTTGGCAAGCAGGCTGTTCGATACTGCTATA 1380  
 DB 1321 GGCATTTTATATATGTTGATTCATCAAGTTTGGCAAGCAGGCTGTTCGATACTGCTATA 1380  
 QY 1381 TCCTGTATTTCTGTTTATTTGTTTATTTCTGAGAAATATGTTAAGATCTCTGCTG 1440  
 DB 1381 TCCTGTATTTCTGTTTATTTGTTTATTTCTGAGAAATATGTTAAGATCTCTGCTG 1440  
 QY 1441 ATTTGGAATTTCTCTATTTCTTCAATTTAAATTTTCTCAAACTCTTCTTCTGCGAAGCAT 1500  
 DB 1441 ATTTGGAATTTCTCTATTTCTTCAATTTAAATTTTCTCAAACTCTTCTTCTGCGAAGCAT 1500  
 QY 1501 TTCTTGTTTACCCAAATCAACCTATTTCTGAAATATGATGGTTAGCAAGTTGAGATA 1560  
 DB 1501 TTCTTGTTTACCCAAATCAACCTATTTCTGAAATATGATGGTTAGCAAGTTGAGATA 1560  
 QY 1561 ACTAGAGCCTGTAATCCATCATTTTAAATGGCAATGATATGACAGTTTATTTTATGTT 1620  
 DB 1561 ACTAGAGCCTGTAATCCATCATTTTAAATGGCAATGATATGACAGTTTATTTTATGTT 1620

QY 1621 ATATATAAACCTCAACAATTTTCCAAACAATTTACCAAAATGGTCATTAACTGTATCCA 1680  
Db 1621 ATATATAAACCTCAACAATTTTCCAAACAATTTACCAAAATGGTCATTAACTGTATCCA 1680  
QY 1681 CAAAGGATTTCTGCATTAACATCTTTAAACAATTTACCTAAATTTAGTGCATATTAA 1740  
Db 1681 CAAAGGATTTCTGCATTAACATCTTTAAACAATTTACCTAAATTTAGTGCATATTAA 1740  
QY 1741 ACTTATTTGGTGGCATGACTATATCAACAGTTGCATGATATATGATACAAATTTGTA 1800  
Db 1741 ACTTATTTGGTGGCATGACTATATCAACAGTTGCATGATATATGATACAAATTTGTA 1800  
QY 1801 TTCTTTTCCATTTGCACTGAAATACCATTAATAAAGAGAAATCCCATCATCAAAATTTGA 1860  
Db 1801 TTCTTTTCCATTTGCACTGAAATACCATTAATAAAGAGAAATCCCATCATCAAAATTTGA 1860  
QY 1861 GCCTATATTGATTGATFACCTCAGAAGAATCTGGCAGTAGGAGCTATAAAGGATTAAGCAA 1920  
Db 1861 GCCTATATTGATTGATFACCTCAGAAGAACTGGCAGTAGGAGCTATAAAGGATTAAGCAA 1920  
QY 1921 TTGGGAAAGGATTTGGGAAGTTGGTAGTGAACATCTTCTCACCTGGAGCTCATGAGCAA 1980  
Db 1921 TTGGGAAAGGATTTGGGAAGTTGGTAGTGAACATCTTCTCACCTGGAGCTCATGAGCAA 1980  
QY 1981 CTTTGAATAGTTTAACTGTGATGCATATGTAGATTTCTAACACATTTTCCCTTGAATA 2040  
Db 1981 CTTTGAATAGTTTAACTGTGATGCATATGTAGATTTCTAACACATTTTCCCTTGAATA 2040  
QY 2041 GAAATTTGGCACAACAATTTTAAATTAATTTAGCAATATTTGGATATTTAAAGCTTCT 2100  
Db 2041 GAAATTTGGCACAACAATTTTAAATTAATTTAGCAATATTTGGATATTTAAAGCTTCT 2100  
QY 2101 TATAGAAAGAGATPACCTGTATATTTAAGCCATGATGAGGTATATACATGTTATATAT 2160  
Db 2101 TATAGAAAGAGATPACCTGTATATTTAAGCCATGATGAGGTATATACATGTTATATAT 2160  
QY 2161 TACTGTACATGCAATTAATTTTATTTTATCATTTGAGTGCATTTCTTTAAATTTAGT 2220  
Db 2161 TACTGTACATGCAATTAATTTTATTTTATCATTTGAGTGCATTTCTTTAAATTTAGT 2220  
QY 2221 AATGCCTTTTGGCTTTAAATTTTCTCCTGATATTTAAATAGATACAGTAACCTTTCAATG 2280  
Db 2221 AATGCCTTTTGGCTTTAAATTTTCTCCTGATATTTAAATAGATACAGTAACCTTTCAATG 2280  
QY 2281 TTAGTGTCTGTAATTTTTCATCTCTATTTTGAACCATTTTATTTCCACATGTG 2340  
Db 2281 TTAGTGTCTGTAATTTTTCATCTCTATTTTGAACCATTTTATTTCCACATGTG 2340  
QY 2341 CTCTTAATAAGTAGCATATAGTTAAATTTTAAAAATCCAAATATGCAATCACCTTTTAG 2400  
Db 2341 CTCTTAATAAGTAGCATATAGTTAAATTTTAAAAATCCAAATATGCAATCACCTTTTAG 2400  
QY 2401 GTTAAAAATTTAATCCATTTACATTTGTGACAAATTCGACATATATATGTTCTFAAATCTA 2460  
Db 2401 GTTAAAAATTTAATCCATTTACATTTGTGACAAATTCGACATATATATGTTCTFAAATCTA 2460  
QY 2461 TCATCTTTACTAGTGGTTTCCATTTCTCTGCTCCAAATATTTTTCACAGCTTTATAA 2520  
Db 2461 TCATCTTTACTAGTGGTTTCCATTTCTCTGCTCCAAATATTTTTCACAGCTTTATAA 2520  
QY 2521 CACAACCTTTTATAGAAAAGTTATACATAACACAGCATCAACTATTTTCAAGAACCCCAAT 2580  
Db 2521 CACAACCTTTTATAGAAAAGTTATACATAACACAGCATCAACTATTTTCAAGAACCCCAAT 2580  
QY 2581 AAGCAACAAAACACAGACTAACAAAATGTGTAAACAAGAACTAATGACCTTTCTAAATC 2640  
Db 2581 AAGCAACAAAACACAGACTAACAAAATGTGTAAACAAGAACTAATGACCTTTCTAAATC 2640  
QY 2641 AAACATTTCAATTCTACATGCTATTTTACAAACAGGGGAAACTCCATGTTTACAGGC 2700  
Db 2641 AAACATTTCAATTCTACATGCTATTTTACAAACAGGGGAAACTCCATGTTTACAGGC 2700

QY 2701 ATGTCATATTGAAATTAAGCTGCAATAGCTTTTATACAAATATGCTCTCAAGAAAT 2760  
Db 2701 ATGTCATATTGAAATTAAGCTGCAATAGCTTTTATACAAATATGCTCTCAAGAAAT 2760  
QY 2761 GAATCATTAAGACAGTAAATTAGGAGTTTCACAAAATTTAAACATTTACGTAATTTTAAAT 2820  
Db 2761 GAATCATTAAGACAGTAAATTAGGAGTTTCACAAAATTTAAACATTTACGTAATTTTAAAT 2820  
QY 2821 TATTGTCTTCAATTAATTTTAAATTTATTGAAGTCTGAGTTTCAAAAGTGATTTTTCAC 2880  
Db 2821 TATTGTCTTCAATTAATTTTAAATTTATTGAAGTCTGAGTTTCAAAAGTGATTTTTCAC 2880  
QY 2881 AAAGTGCCAACTTAAGCTTAGAGCTTTTCACTGTTTAACTTTTGGCCCTTAAAGTTAAGACA 2940  
Db 2881 AAAGTGCCAACTTAAGCTTAGAGCTTTTCACTGTTTAACTTTTGGCCCTTAAAGTTAAGACA 2940  
QY 2941 TATTCTGAGAATCATATATAGTACATGATTTCTGATGCTATCTGTTTAAATAACAAA 3000  
Db 2941 TATTCTGAGAATCATATATAGTACATGATTTCTGATGCTATCTGTTTAAATAACAAA 3000  
QY 3001 GATTTTCACATGAATACCTATGTAACAAATCTCCATGTTCTACACATATATACCCAGAAC 3060  
Db 3001 GATTTTCACATGAATACCTATGTAACAAATCTCCATGTTCTACACATATATACCCAGAAC 3060  
QY 3061 TTAAGTATATAATATAATAAACATAGCAAAAGCCCTTTTAAAAAATAAAAAA 3114  
Db 3061 TTAAGTATATAATATAATAAACATAGCAAAAGCCCTTTTAAAAAATAAAAAA 3114  
RESULT 2  
AAZ07531  
ID AAZ07531 standard; cdna: 850 BP.  
XX AAZ07531;  
AC AAZ07531;  
XX 26-NOV-1999 (first entry)  
DT Human SDCMP3 polypeptide encoding cdna.  
DE Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;  
KW Schering dendritic cell membrane protein; dendritic cell physiology;  
KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
KW cancer; forensic; human; lectin 73; ss.  
XX Homo sapiens.  
OS Key Location/Qualifiers  
FH CDS 108..596  
FT /\*tag= a  
FT /product= "SDCMP3"  
XX WO9947673-A2.  
XX 23-SEP-1999.  
XX 16-MAR-1999; 99WO-US03740.  
XX 17-MAR-1998; 98US-0040111.  
XX (SCHE ) SCHERING CORP.  
XX Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
PI Philipps JH;  
XX WPI: 1999-562114/47.  
DR P-PSDB; AAY27446.  
XX Binding compound specific for primate or rodent Schering dendritic cell  
PT membrane proteins  
PS Claim 8; Page 80; 89pp; English.  
XX The invention relates to a binding compound comprising an antibody  
CC





Db 206 GCTGCTGATTTCCTTCCAGTACCTTCTGAGTACCTGTTTCATTGCGAGCTGTGGTGACT 265  
Qy 280 TACCATTTTTACATATGTTGAAGCTGGAAAAGGCTGTCTGAACCTACACTCATATCATTTCA 339  
Db 266 TACCAATTTTATATGAGACCTGAGAGACTATATGAACCTTACACATACCATTC 325  
Qy 340 AGTCTCACCTGCTTCACTGAGGACAAAGGTGCCAG-----CCTGGGGATCTGCCCA 393  
Db 326 AGTCTCACCTGCTTCACTGAGGAGCTATGTTGCAGAAAAATGTGGGATGCTGCCCA 385  
Qy 394 GCTTCTTGGAGTCAATTTGGTTCAGTTCCTACTTTCATTTCCAGTGAAGAAAGTTTGG 453  
Db 386 AATCACTGGAAGTCATTTGGCTCCAGCTGCTACTCTCATTTTCTACCAAGGAGACTTCTGG 445  
Qy 454 TCTAAGAGTGAAGCAACTGTGTGAGATGGGAGCAGATTTGGTTGTTCACACAGAA 513  
Db 446 AGCACCAGTGAAGCAACTGTGTGAGATGGGAGCTCATCTGTGTGTGATCAACTGAA 505  
Qy 514 GCAGAGCAGAAATTCATTTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 573  
Db 506 GCGAGCAGAAATTCATCCACCAGCAGCTGAATGAGTCAATTTCTTACTTCTGGGTCTT 565  
Qy 574 TCAGACCCAGCAAGTAATAATATGGAATGGATTAAGACACCTTATGAGAAAAAT 633  
Db 566 TCGGATCCACAAGGTAATGCAATGGCAATGGATGATGATCTCTTTCAGTCAAAAT 625  
Qy 634 GTGAGATTTGGCACCTAGCTGAGCCCAATCATCTGCAGAGCAATGTGCTTCAATAGTC 693  
Db 626 GTCAGGTTCTGGCACCCCTCATGAGCCAGCTGAATGAGTCAATTTCTTCAATAGTT 685  
Qy 694 TTCTGGAACCTACAGATGGGCTGGAATGATGTTATCTGTCAAACTAGAAGCAATTC 753  
Db 686 TACTGGAATCCTTGAATGGGCTGGAATGATGTTTCTGTCATAGTAACACAAATTC 745  
Qy 754 ATATGAGATGAATAAGATTTACCTATGAT 785  
Db 746 ATATGTAATGAAGAAGATTTACCTATGAT 777

RESULT 4  
AAS01378  
ID AAS01378 standard; cDNA: 1252 BP.  
AC AAS01378;  
XX  
XX  
DE 04-JUL-2001 (first entry)  
XX  
XX  
DE Murine TANGO 405 alternative splice variant cDNA sequence.  
XX  
XX  
KW Murine; TANGO 210; clone jtmMa025a11; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.  
XX  
XX  
OS Mus sp.  
XX  
XX  
FH Key  
FT 179..808  
FT /\*tag= a  
FT /product= "TANGO 405 alternative splice variant protein"  
XX  
XX  
XX WO200118016-A1.  
XX  
XX  
XX 15-MAR-2001.  
XX  
XX  
XX 30-JUN-2000; 2000WO-US18174.  
XX  
XX  
XX 10-SEP-1999; 99US-0393996.  
XX  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX

PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI; 2001-183280/18.  
DR P-PSDB; AAU00482.  
XX  
PT Isolated nucleic acid molecules encoding proteins useful as modulating  
agents in regulating a variety of cellular processes are used for  
treating e.g. cancer and autoimmune disorders -  
PT  
XX Disclosure: Fig 6N-6P; 326pp; English.  
XX  
CC The present sequence encoding for murine TANGO 405 alternative splice  
variant protein is isolated from cDNA clone jtmMa025a11 from a long-term  
bone marrow cDNA library. TANGO 405 (AAU00480) is 1 of 3 novel murine  
proteins which include TANGO 210 (AAU00470) and INTERCEPT 400  
(AAU00477). Six novel human proteins which include TANGO 210 (AAU00469),  
TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473),  
INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat  
INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids  
encoding these novel proteins are useful as modulating agents in  
regulating a variety of cellular processes and can be used to express  
the proteins in a host cell in gene therapy applications. Human and  
murine TANGO 405 proteins show sequence homology to murine dectin-2.  
TANGO 405 modulates growth, proliferation, survival, differentiation,  
activity, morphology and movement/migration of human lymphocytes and  
bone marrow cells and tissues and can be used to prevent, diagnose or  
treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 1252 BP; 330 A; 283 C; 286 G; 353 T; 0 other;  
Query Match 12.2%; Score 379.2; DB 22; Length 1252;  
Best Local Similarity 73.8%; Pred No. 2.1e-61;  
Matches 555; Conservative 0; Mismatches 178; Indels 19; Gaps 5;  
Qy 40 GCAGTTTCTCCCTGAGCTCTAGCTTCTTTTAAATGAAGCTCTCTGGGCAACATCTTT 99  
Db 72 GGAAGTTGATCTGAACCTCTGGCCCTTTTGACAGAAGCCAGCTCCCTGAGTCGTATTTT- 130  
Qy 100 AGGGAGAGAGTACAAAAGGTTCTTGGACCTTCTCAACACAGGAGCCCTGCATATGATG 159  
Db 131 --GGAGACAGATGCAAGAAACCCCT-GACCTTCTCAACATA---CACCTCAACAATGTTG 184  
Qy 160 CAAGAGCAGCAACCTCAAGAGTACAGAAAAGAGGCTGGTGTCTGCTGAGACTCTGGTCT 219  
Db 185 CAGGAAGACATCCCA-----AGGAGAGGAGTCTCTGCGACCTGAGACTCTGGTCA 238  
Qy 220 GTGGCTGGGATTTCCATTGCACTCTCTCAGTGTCTGCTTCAATTTGTGAGTGTGTAGTAACT 279  
Db 239 GCTGCTGTGATTTCCATGTTTACTCTTGAGTACCTGTTTCAATTCGAGCTGTGTGTGACT 298  
Qy 280 TACCATTTTACATATGTTGCAAACTGGCAAAAGGCTCTGCTGACTACACTCATATCATTTCA 339  
Db 299 TACCATTTTATTTGACACAGCCAGTGAAGACTATATGAACCTTCACACATACCATTCCTCC 358  
Qy 340 AGTCTCACCTGCTTCACTGAGGACAAAGGTGCCAG-----CCTGGGGATGTTGCCCA 393  
Db 359 AGTCTCACCTGCTTCACTGAGGAGCTATGGTGTGAGAAAATGTGGGATGCTGCCCA 418  
Qy 394 GCTTCTTGGAGTCAATTTGGTTCCAGTGTCTACTTTCATTTCCAGTGAAGAGGTTTGG 453  
Db 419 AATCACTGGAAGTCAATTTGGTCTCCAGCTACCTTCATTTCTACCAAGGAACTCTCTGG 478  
Qy 454 TCTAAGAGTGAAGCAACTGTTTCAGATGGGAGCAGATTTGGTGTGTTTCAACACAGAA 513  
Db 479 AGCACCAGTGAAGCAACTGTTTCAGATGGGGCTCATCTGGTGGTGTATCAATACTCAA 538  
Qy 514 GCAGAGCAGAAATTTTCATTTGCTCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 573  
Db 539 GCGGAGCAGAAATTTTCATCACCAGCAGCTGAATGAGTCACTTTCTTACTTCTGGGTCTT 598  
Qy 574 TCAGACCCAGCAAGTAAATAATTTGCAATGGATTTGATAAGACACCTTATGAGAAAAAT 633  
Db 599 TCGGATCCACAAGGTAATGGCAATGGCAATGGATGATGATGATCTCTTCTTCAATCAAAAT 658



PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX WPT; 1999-591920/51.  
 XX  
 PT New nucleic acid sequences expressed in ovarian, and some other, cancer  
 PT tissues, and derived polypeptides, for treatment of ovarian cancer and  
 PT identification of therapeutic agents  
 XX  
 XX  
 PS Claim 3; Page 206-207; 310pp; German.  
 XX  
 CC This invention describes novel nucleic acid (cDNA) sequences (A) which  
 CC have anticancer activity and are highly expressed in ovarian tumor  
 CC tissue (and some also in testis and breast cancer tissue). The products  
 CC of the invention can be used for gene therapy. (A) are used (i) for  
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete  
 CC genes. (B) are used (i) to identify agents suitable for treatment of  
 CC ovarian cancer; (ii) directly for treating this form of cancer  
 CC (including expression from gene therapy vectors) and (iii) for generation  
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed  
 CC sequence tags) from a particular tissue type before comparison of  
 CC expression patterns. This allows a significantly longer fragment of the  
 CC gene to be revealed, so should reduce the number of failures associated  
 CC with the fact that ESTs from different libraries may represent different  
 CC parts of the same unknown gene, distorting the estimated frequency of  
 CC occurrence in a particular tissue. AA277450-277572 represent the human  
 CC ovarian tumor cDNA library derived EST fragments described in the method  
 CC of the invention and encode the protein fragments represented in  
 CC AA277505-Y76638.  
 XX  
 SQ Sequence 3708 BP; 1036 A; 818 C; 671 G; 1183 T; 0 other;

Query Match 11.7%; Score 364.6; DB 20; Length 3708;  
 Best Local Similarity 88.3%; Pred. No. 1.3e-58;  
 Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;  
 QY 2500 TATTTTACAGCTTATACACAACTTTTATAGAAAGTTATACATAACAGCATC 2559  
 DB 40 TTTTATTTTACAGCTTATACACAACTTTTATAGAAAGTTATACATAACAGCATC 99  
 QY 2560 AACTATTCTTCAAG-----AACCCAAATAGCAACAAACAAACAGCACTAACAATGT 2609  
 DB 100 AACTATTCTTCAAGCAACAATTAACCCGATAGCAACAAACAAACAGCACTAACAATGT 159  
 QY 2610 GTAACAAGAACTAATGACCTTTCTAAATCAACAACTCAATATCTACAATGCTATT 2669  
 DB 160 GTAACAAGAACTAATGACCTTTCTAAATCAACAACTCAATATCTACAATGCTATT 219  
 QY 2670 ACAACAGGGAACCTCCATGTTTACAGGCATGTCATATTGAAATAAAGCTCCATAG 2729  
 DB 220 ACAACAGGGAACCTCCATGTTTACAGGCATGTCATATTGAAATAAAGCTCCATAG 279  
 QY 2730 C-TTTTATACAATTATCGCTCTCAAGAAATGAATCAATTAAGCAGCTAATTAGAGTTC 2788  
 DB 280 CAATTTATACAATTATCGCTCTCAAGAAATGAATCAATTAAGCAGCTAATTAGAGTTC 339  
 QY 2789 ACAATTTAAACATTTTCAGCTAATTTTAAATTAATTTGCTTCAATTAATTTAATTTG 2848  
 DB 340 ACAATTTAAACATTTTCAGCTAATTTTAAATTAATTTGCTTCAATTAATTTAATTTG 387  
 QY 2849 AAGTCTGAGTTTCAAAAGTGA---TTTPTTCCCAAAAGGTGCCAACACTTTAAGCTAGAG 2905  
 DB 388 AAGTCTGAGTTTCAAAAGTGAATTTTTCACCAAAAGTTTCAACACTTTAAGCTAGAG 447  
 QY 2906 CTTTCAGTGTTAATCTTGGCCT-AAAAGTTAAGACATATCTGAGATCATTAATAGTCAC 2964  
 DB 448 CTTTCAGTGTTAATCTTGGCCTAAAAGTTAAGACATATCTGAGATCATTAATAGTCAC 507  
 QY 2965 ATGATTTCTGATGCTATCTGCTCTGTTAATAACAAAGATTTTCA 3007  
 DB 508 ATGATTTCTGATGCTATCTGCTGTTAATAACAAAGCTTTTA 550

RESULT 7

AAC77686/C  
 ID AAC77686 standard; cDNA; 3748 BP.  
 XX  
 AC AAC77686;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated gene sequence SEQ ID NO:80.  
 XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;  
 KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 KW dermatologic; neuropathic; thrombotic; coagulant; neuropathic;  
 KW vasotonic; antipsoriatic; antidiabetic; antidiabetic; antidiabetic;  
 KW immune disorder; haematopoietic cell disorder; gene therapy; inflammation;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombotic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO200055350-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05882.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2000-587533/55.  
 DR P-PSDB; AAB43477.  
 XX  
 PT Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer -  
 XX  
 PS Claim 1; Page 671-673; 2352pp; English.  
 XX  
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnary; immunomodulator;  
 CC antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 CC antiinflammatory; antithyroid; antidiabetic; antidiabetic; antidiabetic;  
 CC dermatologic; neuropathic; thrombotic; coagulant; neuropathic;  
 CC vasotonic; antipsoriatic; antidiabetic; antidiabetic; antidiabetic;  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombotic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.  
 XX  
 SQ Sequence 3748 BP; 1199 A; 685 C; 825 G; 1039 T; 0 other;

Query Match 11.7%; Score 364.6; DB 21; Length 3748;  
 Best Local Similarity 88.3%; Pred. No. 1.3e-58;  
 Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;  
 QY 2500 TATTTTACAGCTTATACACAACTTTTATAGAAAGTTATACATAACAGCATC 2559  
 DB 3725 TTTTATTTTACAGCTTATACACAACTTTTATAGAAAGTTATACATAACAGCATC 3666



Qy	2916	TAACTTGGCCCT-AAAAGTTAAGACATAATCTCAGAATCAATAATAGTCACATGATTTCTG	2974
Db	3239	TAACTTGGCCCTAAAAAGTTAAGACATAATCTCAGAATCAATAATAGTCACATGATTTCTG	3180
Qy	2975	ATGCTATCTCTCTGTTTAATACAAAGATTCA	3007
Db	3179	ATGCTATCTGTCGTGTTAATAAAGTCTTTA	3147
RESULT 9			
AAI97506			
ID	AAI97506	standard; cDNA; 803 BP.	
XX	AAI97506;		
XX	13-NOV-2001	(first entry)	
XX	Human neuroblastoma	expressed polynucleotide SEQ ID NO 3581.	
XX	Human; neuroblastoma;	malignancy; cancer; tumour marker; N-myc; TrkA; ss.	
XX	Homo sapiens.		
OS	WO200166719-A1.		
XX	13-SEP-2001.		
XX	02-MAR-2001;	2001WO-JP01629.	
XX	07-MAR-2000;	2000JP-0159195.	
XX	(CHIB-) CHIBA PREFECTURE.		
PA	(HISM ) HISAMITSU PHARM CO LTD.		
PI	Nakagawara A;		
WPI	2001-565584/63.		
Nucleic acids originating in gene expressed in human neuroblastoma,			
useful as probe or primer in diagnosing prognosis of human			
neuroblastoma, malignancy and susceptibility indicator or tumour marker			
for anti-cancer agents -			
Claim 1; Page 2600; 2979pp; Japanese.			
The invention relates to novel genes (AAI93926-AAI97963) expressed in			
human neuroblastoma. The nucleic acids are applicable as a probe or			
primer in diagnosing the prognosis of human neuroblastoma, malignancy and			
susceptibility indicators or tumour markers for anti-cancer agents. The			
gene information for diagnosing prognosis is related to factors similar			
to that for N-myc and TrkA genes.			
Sequence 803 BP; 285 A; 137 C; 112 G; 251 T; 18 other;			
Query Match	11.3%;	Score 351.4; DB 22; Length 803;	
Best Local Similarity	86.9%;	Pred. No. 2.9e-56;	
Matches 433; Conservative	0; Mismatches 41; Indels 27; Gaps 5		
Qy	2500	TATTTTTCACGCTTATAACACAACTTTTATTAGAAAAGTTATACATAACACAGCATC	2559
Db	26	TTTTTTTTCACGCTTATAACACAACTTTTATTAGAAAAGTTATACATAACAGCATC	85
Qy	2560	AACATTTTTCAG- - - - - AACCCAATAAGCAACAAAACCAACACTAACAAATGT	2609
Db	86	AACTATTTTTCAGACAATATTTAAACCCGATTAAGCAACAAAACCAACACTAACAAATGT	145
Qy	2610	GTAAACAAGAACTAATGACCTTCTTAAATCAAACTCAATTTATCTACAAATGCTATTT	2669
Db	146	GTAAACAAGAACTAATGACCTTCTTAAATCAAACTCAATTTATCTACAAATGCTATTT	205
Qy	2670	ACAAACAGGAAAACCTCCATGGTTTACAGGCATGTCATTTTGAATAAAGCTGCAATAG	2729

An alternative splice variant of this protein is also described (AAU000482). Other novel murine proteins include TANGO 210 (AAU000470) and INTERCEPT 400 (AAU00477). Six novel human proteins which include TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids encoding these novel proteins are useful as modulating agents in regulating a variety of cellular processes and can be used to express the proteins in a host cell in gene therapy applications. Human and murine TANGO 405 proteins show sequence homology to murine Dectin-2. TANGO 405 modulates growth, proliferation, survival, differentiation, activity, morphology and movement/migration of human lymphocytes and bone marrow cells and tissues and can be used to prevent, diagnose or treat leukaemia, lymphomas and autoimmune disorders.

Sequence 821 BP: 214 A; 188 C; 195 G; 223 T; 1 other;

QY	2504	TTTTTTACAGCTTATAACACAACCTTTTATTAGAAAAGTTATTACATAACACAGCATCAACT	2563
Db	24425	TTTTTTTCAACTTATAACACAACCTTTTATTAGAAAAGTTATTACATAACATAGCATCAACT	24366
QY	2564	ATTTTCAAGAAC-----CCAATANGCAACAAAACGACATTAACAAAATGTGTAAC	2614
Db	24365	ATTTTCAAGACATATTAAACCCATAGCAAAAACACAGCTTACAAAATGTGTAAC	24306
QY	2615	AGAAACCTAATGAOCTTTCTAAAATCAACACATTCAATATTCTACAAATGTCTATTATTACAAA	2674
Db	24305	AGAAACCAATGAGCTTTCTAAAATCAACACATTCAATATTCTACAAATGTCTTTTATTACAAA	24246
QY	2675	CAGGGAAACCTCCATGTTTACAGGCATGTCATTGAAAAATAAGCTGCAATAGCTTTT	2734
Db	24245	CAGGGAAACCTCCATGTTTACAGGCACGATATTGAAAAATAATCTGCAATAGCAATT	24186
QY	2735	TATACAATTATGCTCTCAAGAAAATGAATCATTAAGACAGAGTAAATGAGAGTTCACAAAT	2794
Db	24185	TTTTATGGTTAAACTCTCAAGAAACGGAATCATTAACACAGTAAT--GAGTTCACAAAT	24129
QY	2795	TTAAACATTTACAGTAAATTTAAATATTGTCTTCAATAATTTTAAATATTGAGTCT	2854
Db	24128	TTAAACATTTTACATAATTTTAAATATTGAGT-----ATACACTGAGTCT	24081
QY	2855	GAGTTTCAAAGTGATTTTTTCCCAACAAAGTGCCACACACTTAAGCTAGAGCTTTTCAGTG	2914
Db	24080	GTTTCAAAGTGATTTTTTTCCTCCACGAAAGTGTCACACACTTAAGCTAGAACTTTTCAGTG	24021
QY	2915	TTAACTTGGCCCT-AAAAGTTTAACACATATTTCTGAGATCAATAATAGTACACATGATTTCT	2973
Db	24020	TTAATTTTGGCCCTAAAAGTTTAAAGACATATCTGATAAATCAATAAGTACATGATTTCT	23961
QY	2974	GATGCTATCTGCTGTGTTAATA	2995
Db	23960	GATGCTATCTGCTGTGTTAATA	23939

RESULT 12

AAA35003/c  
ID AAA35003 standard; DNA; 56583 BP.

AA  
AC AAA35003;

28-JUL-2000 (first entry)

Human adenosine receptor related polynucleotide SEQ ID NO:2692.

xx Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX PN WO200009525-A2.

XX  
PD  
24-FEB-2000.

XX 03-AUG-1999: 99W0-IIS17712.

XX 03-AUG-1998. 98US-0095212

XX  
DA (JFEC - ) UNIV EAST CAROLINA

XX	W-20	703
DT		

XX 300 305 371 110

XX  
XX

PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
PT cancers -

PS Disclosure; page 903-916; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, anti-allergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA33233 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 56583 BP; 16610 A; 11169 C; 11927 G; 16877 T; 0 other;  
XX SO

Query Match 10.2%; Score 317.6; DB 21; Length 56583;

Query Match 10.2%; score 317.0; ED 2  
Best Local Similarity 83.3%; Pred. No. 1.1e-49;

BEST LOCAL SIMILARITY	83.38,	FILED: NO. 1-10-45,	
MATCHES	418;	CONSERVATIVE	0;
MISMATCHES	0;	MISMATCHES	59;
INDELS	25;	INDELS	25;
GAPS	4;	GAPS	4;

QY 2504 TTTTTCACGCTTATAACACAACCTTTTATTAGAAAAGTTATACATATACACAGCATCAACT 2563

QY 2564 ATTTTCAAGAC-----CCAAATAGCAACAAACCAGACTACAAAAATGTGTAACT 2614

QV 2615 AAGAACTAATGACCTTTCTAAATCAACATTTCRAATTATCTACAATGTCTATTACAA 2674

[illegible]

DD	24 245	CAGGGAAACATCCATGGTTATACAGGCACACACGATATGGAAATATATATATCTGCATATAGCAAT	24 101
OV	2735	TATACATTTATCGCTCTCAAGAAATATGATTCATTTAGACACGCTAATTAGGAGTTCACAAAT	2794

DB 24185 TTTTATGTTAAACTCTCAAGAAGCGGAACATTAAAACAGTAAT---GAGTTCACACAAAT 24191  
2705 TTATAAACAATATTCCTCAAATAAATTTTAAATTTATCAGAGTCT 2854

Db 24128 TTTAAACATTTTCACATAATTTTAAATTTATTTGAGT-----ATACACGAAAGTCT 2409

Db 24080 GTTTCAAAGTGATTTTTTCCCCACGAAAGTGTCAACACTTAAAGCTAGAACTTTTCAGTG 24073

Db 24020 TTAATTTGGCCCTAAAAGTTAAGACA

Qy 2974 GATGCTATCTGCTCTGTTAATA 2995



## RESULT 13

AAV42554  
ID AAV42554 standard; cDNA; 501 BP.

XX AAV42554;  
XX

DT 09-NOV-1998 (first entry)  
XX

DE Mouse dectin-2 extracellular domain cDNA.

XX Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
KW drug screening; ss.

XX Mus sp.  
XX

PN W09828332-A2.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-US23761.

XX 20-DEC-1996; 96US-0772440.

XX (TEXA ) UNIV TEXAS SYSTEM.

PA Ariizumi K, Takashima A;

XX WPI; 1998-377594/32.

DR P-PSDB; AAW63022.

XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
PT and -2 - useful, e.g. to regulate immune response, as vaccine  
PT adjuvants, for diagnosis and drug screening

XX Disclosure; Page 153; 200pp; English.

XX This cDNA sequence codes for the extracellular domain (see  
CC AAW63022) of a novel murine dendritic cell (DC) specific member of  
CC the C-type lectin family, termed DC-associated C-type lectin-2, or  
CC dectin-2 (see also AAW63010), that is essential for DC-mediated T  
CC cell activation. Full-length cDNA (see AAV42549) was isolated by  
CC subtractive hybridisation from BALB/c mouse epidermis DC line XS52.  
CC The invention provides: dectin-1 and -2 polypeptides (see also  
CC AAW63009-22 and AAW69236-37), useful for purifying T cells, for  
CC detecting autoantibodies and for up-regulating immunity e.g. as  
CC vaccine adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and  
CC AAV44850-51); expression vectors; recombinant host cells; probes  
CC and primers; antibodies; compounds that modulate dectin-mediated  
CC activation of T cells; transgenic animals; and dectin ligands.  
CC Dectin expression can also be down-regulated to treat allergy and  
CC autoimmune disease. Recombinant extracellular domain has been  
CC expressed as a His-tagged protein (see AAW69237) useful for antibody  
CC production.

XX Sequence 501 BP; 142 A; 108 C; 115 G; 136 T; 0 other;

Query Match 9.7%; Score 303.2; DB 19; Length 501;

Best Local Similarity 78.4%; Pred. No. 2.4e-47;

Matches 378; Conservative 0; Mismatches 98; Indels 6; Gaps 1;

QY 305 GCAAAAGCTGTCTGAACATACACTATATCATATCAAGTCTCACCTGCTTCAGTGAAGGGA 364

DB 20 GTAGAAGACTATATGAACATTCACATACCATTCAGTCTCACCTGCTTCAGTGAAGGGA 79

QY 365 CAAAGGTGCCAG-----CCTGGGATCTGCCAGCTCTTGGAGTCATCTTGGTTCCA 418

DB 80 CTATGGTCTCAGAAAATATGTGGGATGCTGCCAAATCACTGGAAGTCATTTGGCTCCA 139

QY 419 GTTGCTACTTCATTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTG 478

Db 140 GCTGCTACCTCATTTTACCAGGAGAACTTCTGGAGCACCAGTGAGCAGAACTGTGTT 199  
QY 479 AGATGGGAGCACATTTGGTGTTCACACAGACAGACAGCAATTTTATTGTCAGC 538  
Db 200 AGATGGGAGCTCATCTGGTGTGATCAATCTGAAGGGAGCAGAAATTCATCACCAGC 259  
QY 539 AGCTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCCAAGGTAATAAAT 598  
Db 260 AGCTGAATGAGTCATTTTCTTACTTCTGGGTCTTCGGATCCACAAAGGTAATGGCAAT 319  
QY 599 GGCAATGGAITGATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGC 658  
Db 320 GGCAATGGAITGATGATGATCTTTTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAAC 379  
QY 659 CCAATCAATCTGCAGAGCAATGTCCTTCAATAGTCTTCTGAAACCTACAGGATGGGGCT 718  
Db 380 CCAATCTTCAGAGAGCGGTGTTTCAATAGTTTACTGGAATCCCTCGAAATGGGGCT 439  
QY 719 GGAATGATGTTATCTGAAACTAGAGGAATTCATATATGTGAGATGAATAAGATTACC 778  
Db 440 GGAATGATGTTTCTGTGTAGTAACACAAATTCATATGTGAAATGAAGAAGATTACC 499  
QY 779 TA 780  
Db 500 TA 501

## RESULT 14

AAF90241

ID AAF90241 standard; cDNA; 1312 BP.

XX AAF90241;

XX 06-AUG-2001 (first entry)

DE Nucleotide sequence of BDCA-2 antigen.

XX BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;

XX viral infection; autoimmune disease; allergic response; cancer; ss.

XX Homo sapiens.

Key Location/Qualifiers

CDS 491..1132

FT /\*tag= a

FT /product= "BDCA-2"

XX W0200136487-A2.

XX 25-MAY-2001.

XX 15-NOV-2000; 2000WO-IB01832.

XX 15-NOV-1999; 99US-0165555.

XX 23-NOV-1999; 99US-0167076.

XX 28-JAN-2000; 2000US-0179003.

XX 07-FEB-2000; 2000US-0180775.

XX 11-APR-2000; 2000US-0196824.

XX 13-APR-2000; 2000US-0197205.

XX (MILT-) MILTENYI BIOTECH GMBH.

XX Schmitz J, Dzionek A, Buck DW;

XX WPI; 2001-355622/37.

XX P-PSDB; AAB84215.

XX Compositions and cell populations enriched in dendritic cells through  
PT use of antigen-binding fragments specific for BDCA-2, BDCA-3 or  
PT BDCA-4, are used to treat viral infections, autoimmune diseases,  
PT allergic responses and cancer -  
XX Claim 134; Fig 12; 115pp; English.



QY	223	GTGGGATTTCCATTGGCACTCCTCACTGCTTGCCTTCATTGTGAGCTGTGAGTAACATTAC	282
Db	88	GCAGTCGATCCATCTTGCTCCCTCAGTGTCTGTTTCACTGTGAGTCTTGTTGGTGCCTCAC	147
QY	283	CATTTTACATATGGTGAACCTGCCAAAGCGTCTGAACATA-----CACTCATAT	333
Db	148	AAATTATGTATAGCAAACTGTCAAGAGCGTGTCCAAGTACGAGAGTATCAACAGTAT	207
QY	334	CATTCAAGTCTCACTGTTCACTGAAGGACAAAGTGCACGCTCGGGATGTTCGCCA	393
Db	208	CATCCAGCCTGACCTGGTTCATGGGAAGAAAGGACATAGAGATTGSACTGCTGCCCA	267
QY	394	GCTTCTTGGAAAGTCATTTGGTTCAGTGTCTACTTCAATTCAGTGAAGAGAAGTTTGG	453
Db	268	ACCCCTTGGACTTCATTTCACTAGTGTCTACTTTTATTTCTACTGGGATGCAATCTGG	327
QY	454	TCTAAGATGAGCAAGACTGTGTGAGATGGGAGCACATTTTGGTGTCTTCAACACAGAA	513
Db	328	ACTAAGACTCAAAAGAACTGTCTGTGATGGGGGCTGATCTGGTGGTATCAACACACCG	387
QY	514	GCAGAGCAGAAATTCATTTGTCACAGCAGCTGAATGAGTCATTTCTCTATTTTCTGGGGCTT	573
Db	388	GAAGACAGGATTCATCATTCAGATCTGAAGAAGAAATTCCTCTATTTCTGGGGCTG	447
QY	574	TCAGACCCACAAGGTAATAAATTTGCCAATGGATGATGAAGACACACCTTATGAGAAAAT	633
Db	448	TCAGATCCAGGGGGTCGCGGACATTTGCCAATGGTTGACCAACACCATCAATGAAAT	507
QY	634	GTCAAGATTTGGCACTTAGTGAGCCCAATCATTTGCAAGAGCAATGTGCTTCAATAGTC	693
Db	508	GTCAATTTCTGGCACTCAGGTGAACCCATAACCTTGATGAGCGTGTGGCGAATAAAT	567
QY	694	TTCTGGAAACCTACAGGATGGGCTGGAATGATGTTATCTGTGAACCTAGAAGGAATCA	753
Db	568	TTCCGTCTCTCAGAGAATGGGCTGGAATGACATTCACCTGTCTGATACCTCAGAACTCA	627
QY	754	ATATGTGAGATGAATAGATTTTACCTATGAGTAGAAGCTT	793
Db	628	ATTTCGAAGATGAAGAATCTACATATAAATGAATATT	667

Search completed: February 17, 2003, 09:02:36  
Job time : 696.327 secs



```
QY 100 AGGAGAGAGGTACAAAAGGTTCTGGACCTTCTCAACACAGGGAGCCTGCATAATGATG 159
Db 98 --GGAGACAGATGCAAGAAACCCCT--GACCTTCTGAACATA---CACCTCAACAATGGTG 151
QY 160 CAAGAGCAGCAACCTCAAAAGTACAGAGAAAGAGGCTGTTGCTCCCTCAGACTCTGGTCT 219
Db 152 CAGGAAGACAATCCA-----AGGGAAGGAGTCTCTGGACCTCAGACTCTGGTCA 205
QY 220 GTGCTGGGATTTCCATTGSCACTCCTCAGTGTCTTCATTGTGAGCTGTGTAGTAAT 279
Db 206 GCTGCTGTGATTTCCATGTTACTTGTAGTACCTGCTTCATTGCGAGCTGTGTGTGACT 265
QY 280 TACCATTTCATATATGGTGAACACTGGCAAGAGGCTGTCTGAACCTACACTCATATCA 339
Db 266 TACCAATTTATTATGGACACCCAGTATAGAGACTATATGAACCTTCACACATACCATCC 325
QY 340 AGTCTCACCTGCTTCAGTGAAGGCAAAAGTGCCAG-----CCTGGGGATTTGCCCA 393
Db 326 AGTCTCACCTGCTTCAGTGAAGGACTATGGTTCAGAAAAAATGCGGGATGCTGCCCA 385
QY 394 GCTTCTTGAAGTCAATTTGGTTCAGTGTCTTCTTCATTTCAGTGAAGAGGTTTGG 453
Db 386 AATCACTGGAAGTCAATTTGGTTCAGTGTCTTCTTCATTTCAGTGAAGAGGTTTGG 445
QY 454 TCTAAGAGTGACAGCAACTGTCTGAGATGGGAGCACATTTGGTGTCTTCAACACAGAA 513
Db 446 AGCAGAGTGACAGCAACTGTCTGAGATGGGAGTGTCTGAGTGTCTGAGTGTCTGAGT 505
QY 514 GCAGAGCAGCAATTTCAATTTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 573
Db 506 GCGGAGCAGCAATTTCAATTTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 565
QY 574 TCAGACCCACAAGGTAATAAATTTGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 633
Db 566 TCGGATTCACAAGGTAATAAATTTGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 625
QY 634 GTCAGATTTTGGCAGCTAGGTGAGCCCAATCATTTCTGAGAGCAATGCTTCAATAGTC 693
Db 626 GTCAGTTCTGCGACCCCAATGAGCCCAATCATTTCTGAGAGCAATGCTTCAATAGTC 685
QY 694 TTCTGGAACCTTACAGGATGGGCTGGAATGATGTTATCTGTAACCTAGAGGAATCA 753
Db 686 TACTGGAATCCCTCGAAATGGGCTGGAATGATGTTTCTGATAGTAAACACAATCA 745
QY 754 ATATGTGAGATCAATAAGATTTACCTATGAGT 785
Db 746 ATATGTGAAATGAAGAATTTACCTATGAGT 777
```

RESULT 2

```
US-08-772-440-20
; Sequence 20, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

RESULT 3

```
US-08-772-440-20
; Sequence 22, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
```

```
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-20

Query Match 9.7%; Score 303.2; DB 3; Length 501;
Best Local Similarity 78.4%; Pred. No. 7.8e-58;
Matches 378; Conservative 0; Mismatches 98; Indels 6; Gaps 1;
```

```
QY 305 GCAAAAGGCTGCTGAACTACACTCATATCATTCAGTCTCACTGCTTCAGTGAAGGGA 364
Db 20 GTGAAGACTATATGAACCTTCACACATACCATTCACCTCTCAGTGAAGGGA 79
QY 365 CAAAGGTGCCAG-----CCTGGGGATTTGCCAGCTTCTTGGAAAGTCAATTTGGTTCCA 418
Db 80 CTATGGTGTCCAGAAAAATGCGGGATGCTGCCCAATCACTGGAAGTCAATTTGGTTCCA 139
QY 419 GTTGTCTTCAATTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAAGAGTGTGTG 478
Db 140 GCTGCTACCTTCATTTCTACCAAGGAGAACTTCTGGAGCACCAGTGAAGAGTGTGTTC 199
QY 479 AGATGGGAGCACATTTGGTGTCTCAACACAGACAGCAGAGAAATTTCAATTTGTCAGC 538
Db 200 AGATGGGGCTCATCTGGTGTGATCAATCACTGAAGCGGAGAGAAATTCATCACCACG 259
QY 539 AGCTGAATGAGTCAATTTCTTATTTCTGGGGCTTTTCAGACCCACAGTAATAATAAT 598
Db 260 AGCTGAATGAGTCAATTTCTTACTTCTGGGTCTTTCGGATCCACAAAGTGAATGCAAT 319
QY 599 GCAATGGATTTGATAGACACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGC 658
Db 320 GCAATGGATTCGATGATGATCTTTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAAC 379
QY 659 CCAATCATCTGCAGAGCAATGCTTCAATAGTCTTCTTGGAAACCTACAGATGGGGCT 718
Db 380 CCAATCTTCCAGAGAGCGGTGTGTTCAATAGTCTTCTTGGAAATCCTTCGAAATGGGGCT 439
QY 719 GGAATGATGTTATCTGTGAACCTAGAGGAATTCATATGTCAGATGAATAAGATTTACC 778
Db 440 GGAATGATGTTTCTGTGATAGTAAACACAATTCATATGTAATGTAAGAGATTTACC 499
QY 779 TA 780
Db 500 TA 501
```

CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/772,440  
FILING DATE: CONCURRENTLY HERewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UXD:493  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-772-440-22

Query Match 8.4%; Score 260.2; DB 3; Length 393;  
Best Local Similarity 78.9%; Pred. No. 2e-48;  
Matches 310; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 388 TGCCACGCTTCTGGAAGTCATTTGGTTCAGTGTGCTACTTTCATTTCCAGTGAAGAG 447  
DB 1 TGCCAAATCACTGGAAGTCATTTGGCTCCAGCTGTACCTCATTTCTACCAAGGAGAC 60  
QY 448 GTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTTCAC 507  
DB 61 TTCTGGAGCAGGAGTGAAGTGTGTTGAGATGGGAGCACATTTGGTTGTTCAC 120  
QY 508 ACAGAAGCAGACAGAAATTTTCATTTGCCAGCAGTGAATGAGTCATTTTCATTTCTG 567  
DB 121 ACTGAAGCGGAGCAGAAATTTTCATCACCCAGCAGCTGAATGAGTCATTTCTTACTCCTG 180  
QY 568 GGGCTTTTCAGACCCACAGAGTAAATTAATGCGCAATGGATGATTAAGACACCTTATGAG 627  
DB 181 GGTCTTTTCGAGTCCACAAAGGTAATGGCAATGGCAATGGATGATGATGATGATGATGAT 240  
QY 628 AAAAATGTCAGATTTTGGCACCTAGGTGAGGCCCAATCATTTCTGCAGAGCAATGTGCTTCA 687  
DB 241 CAAAATGTCAGTTCTGGCACCCCATGAAACCCCAATCTTCCAGAGAGCGGTGTGTTCA 300  
QY 688 ATAGTCTTCGGAACCTACAGGATGGGCTGGAATGATGATTTATCTGTGAAACTAGAAAG 747  
DB 301 ATAGTTTACTGGAATCTCTGGAATGGGCTGGAATGATGATTTTCTGTGTAGTAAACAC 360  
QY 748 AATTCAATATGTGAGATCAATAAGATTACCTA 780  
DB 361 AATTCAATATGTGAATGAAGAAGATTACCTA 393

RESULT 4  
US-09-111-470-1  
Sequence 1, Application US/09111470  
Patent No. 6277959  
GENERAL INFORMATION:  
APPLICANT: Valladeau, Jenny  
APPLICANT: Ravel, Odile  
APPLICANT: Bates, Elizabeth E.M.  
APPLICANT: Ford, John  
APPLICANT: Saeland, Sem  
APPLICANT: Lebecque, Serge J.E.  
TITLE OF INVENTION: Mammalian Membrane Protein Genes;

TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,470  
FILING DATE: 08-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/053,080  
FILING DATE: 09-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: SF0695  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)852-9196  
TELEFAX: (650)496-1200  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1104 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 242..952  
US-09-111-470-1

Query Match 6.3%; Score 197; DB 4; Length 1104;  
Best Local Similarity 67.8%; Pred. No. 2e-34;  
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGTGCCACGCTTGGGATGTTGCCAGCTTCTTGAAGTCATTTGGTTCAGTGTCTACT 427  
DB 537 AAGACAGACGCTGGAGCTGTTGCCAAAGAAATTTGAAGTCATTTAGTTCCAACTGCTACT 596  
QY 428 TCATTTCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTTGAGATGGGAG 487  
DB 597 TTATTTCATCAATCAGCATCTTGCAAGACAGTGAAGGACTGTGCTAGAATGGAGG 656  
QY 488 CACATTGTTGTTTCAACAGACAGAGCAGAAATTTTCATTTGTCAGCAGCTGAATG 547  
DB 657 CTCACCTGCTGTGATAAAACACTCAAGAAGAGCAGGATTTTCATCTTCCAGAATCTCAAG 716  
QY 548 AGTCATTTCCTTATTTTCTGGGGCTTTCAGACCCCAAGGTAATAATAATTCGCAATGCA 607  
DB 717 AAGATCTGCTTATTTTGTGGGCTCTCAGATCCAGAGGTCAGGCACATTTGGCAATGGG 776  
QY 608 TTGATAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGGCCCAATCAT 667  
DB 777 TTGATAGACACCATACATCAATGAAGTTCCACATTTCTGGCATCCAGTGAGCCAGTGATC 836  
QY 668 CTGCAGAGCAATGTGCTTCAATAGTCTTCTCTGGAA --- ACCTACAGATGGGCTGGAATG 724  
DB 837 CCAATGAGCGCTGCGTTGTGCTAAATTTTCGTAATAACACCCAAAAGATGGGCTGGAATG 896  
QY 725 ATGTTATCTGTGAACACTAGAAGCAATTCATATGTGAGATGAATAGATTTTACCTATGAG 784  
DB 897 ATGTTAAATGTTCTTGGTCTCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACTTATGAA 956  
QY 785 TAGAAGCTT 793

```

Db 957 CTGAACATT 965
||| ||
RESULT 5
US-09-111-470-7
; Sequence 7, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111.470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 279..992
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1348
; OTHER INFORMATION: /note= "poly-A addition motif"
US-09-111-470-7

Query Match 4.4%; Score 137.4; DB 4; Length 1418;
Best Local Similarity 61.6%; Pred. No. 2.5e-21;
Matches 257; Conservative 0; Mismatches 151; Indels 9; Gaps 2;

QY 375 ACCCTGGGAGCTTCCAGCTTCCTTGAAGTCAATTTGGTCCAGCTGCTACTTCATTTC 434
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 AGCTGGAGCTGTGCCAAAGAGTTGAGGCTATTGGTTCCTCCAGCTGCTACTTGGTTC 643
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 435 CAGTGA-----AGAGAAAGTTTGGTCTTAAGAGTGAGCAGACAGCTGTGTGAGATGGGAGC 488
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 CACAGTTCTTCATCAGCATCTTGAACACAGAGTGAGGAGAACTGCTCCCACTGGGTGC 703
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 ACATTGGTTGTTCACACAGAGAGCAGAGAAATTTCAATTGTCAGCAGCTGAATGA 548
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 704 TCATCTAGTGGTGATCCAAAGCCAGAGCAGGAGGATTTTCATCCTGGGATCTTGGACAC 763
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 GTCAATTTTCTATTCTGGGCTTTTCAGACCCACACAGGTAATAATAATTGGCAATGGAT 608
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 764 TCATGCTGCTATTATTATAGGGTTGTGGGATAC---AGCCCATCGCAATGCAATGGT 820
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 TGATAAGACACCTTATGAGAAAAATCTCAGATTTTGGCACCTAGGTGAGCCCAATCATTC 668
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TGATCAGACACCATATGAAGAAAGTATCACATTTCTGGCAAAATGGTGTGAGCCCGAGTGG 880
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 TGCAGAGCAATGTGCTTCAATAGTCTTCTGGAAACCTACAGGATGGGCTGGAATGATGT 728
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 CAATGAAAAATGTGTACAAATAATTTACCGTTGGAAGAGCTGGATGGGCTGGAACGATAT 940
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 TATCTGTGAAACTMAGAAAGTAATCAATATGTGAGATGAATAAGATTACCTATGAGT 785
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 CTCCTGCAGTCTTAAACAGACAGTCAGTTTGTTCAGATGAAGAAAATAAACTTATGAT 997
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-772-440-42
; Sequence 42, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Arizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-772-440-42

Query Match 3.1%; Score 97.6; DB 3; Length 145;
Best Local Similarity 79.9%; Pred. No. 6.8e-13;
Matches 115; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 640 TTTTGGACCTTAGGTGAGCCCAATCAATCTTCGAGACCAATGTGCTTCAATAGTCTTCGG 699
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TTCTGGCACCCCATGAACCCCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTTACTGG 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 AAACCTACAGGATGGGCTGGGAATGATGTATCTGTGAACACTAGAAGGAATTCATATGCT 759
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 AATCCCTCGAATGGGCTGGGAATGATGTTTCTGTGATAGTAACACAATTCATATGCT 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





[illegible]







Search completed: February 17, 2003, 12:10:31  
Job time : 1394.57 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	364.6	11.7	3748	10	US-09-925-301-80	Sequence
C 2	360.6	11.6	693	10	US-09-833-381-994	Sequence
C 3	311.6	10.0	528	10	US-09-833-381-995	Sequence
C 4	274.8	8.8	642	12	US-10-090-466-1	Sequence
C 5	232.4	7.5	549	12	US-10-090-466-3	Sequence
C 6	197.6	6.3	1091	10	US-09-764-870-199	Sequence
C 7	197.6	6.3	1096	10	US-09-764-870-38	Sequence
C 8	197.6	6.3	1104	9	US-09-962-802-1	Sequence 1
C 9	185.2	5.9	1036	10	US-09-833-381-119	Sequence
C 10	153.2	4.9	291	10	US-09-833-381-996	Sequence
C 11	137.4	4.4	1418	9	US-09-862-802-7	Sequence 7
C 12	120.6	3.9	758	10	US-09-833-381-1339	Sequence
C 13	117.4	3.8	968	9	US-09-965-529-40	Sequence 4
C 14	117.4	3.8	997	9	US-09-905-291A-376	Sequence 3
C 15	117.4	3.8	997	9	US-09-802-853-376	Sequence 3
C 16	117.4	3.8	997	9	US-09-907-824-376	Sequence 3
C 17	117.4	3.8	997	9	US-09-907-841-376	Sequence 3
C 18	117.4	3.8	997	9	US-09-304-011-376	Sequence 3
C 19	117.4	3.8	997	9	US-10-174-590-23	Sequence 2



```
QY 2730 C-TTTTATACAAATATCGCTCTCAAGAAAATGAATCAATTAAGACAGTAATAGGAGTTC 2788
Db 3485 CAATTTTATACAAATTAACCACTCTGAAGAACTGAATCAATTAACAGTAATTAACGAGTTC 3426
QY 2789 ACAAAATTTAAACATTTTCAGTAAATTTTAAATTTATGTCITCAATAATTTTAAATTTATTG 2848
Db 3425 ACAAAATTTAAACATTTTCACATAATTTTAAATTTATGSGGT-----ATACACTG 3378
QY 2849 AAGTCTGAGTTTCAAAAGTGA---TTTTTCCCAAAAGGTGCCAACACTTTAAAGCTAGAG 2905
Db 3377 AAGTCTGAGTTTCAAAAGTGATTTTTTTTCCCAAAAGTTTCAACACTTTAAAGCTAGAA 3318
QY 2906 CTTTCAGTGTAACTTTGCCCT-AAAAGTTAAGACATATTCGAGAAATCAATAATAGTCAAC 2964
Db 3317 CTTTCAGTGTAACTTTGCCCTTAAAGCTTAAAGACATATTCGATTAATCAATAACAGTCAAC 3258
QY 2965 ATGATTTCTGATGCTATCTGCTCTTTTAAACAAAGATTTCA 3007
Db 3257 ATGATTTCTGATGCTATCTGCTCTTTAATAAAGTCTTTTA 3215

RESULT 2
US-09-833-381-994/c
; Sequence 994, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 994
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(693)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-994

Query Match 11.6%; Score 360.6; DB 10; Length 693;
Best Local Similarity 76.7%; Pred. No. 2.7e-58;
Matches 468; Conservative 0; Mismatches 135; Indels 7; Gaps 2;

QY 183 AGAGAAAAGAGGCTGGTGTCTCCCTGAGACTCTGGTCTGTGGCTGGGATTTCCATTTGCAC 242
Db 690 AGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCTGGTCTGGTGTGATTTCCATTTTACT 631
QY 243 CCTCAGTGTCTTCATTTGCTGAGCTGTGTAGTAATTAACATTTACATATGTTGAAC 302
Db 630 CTTGAGTACGTGTTTCATTTGCGAGCTGTGTGTGATTTACCAATTTATATGACACGCC 571
QY 303 TGSCAAAAGCTGTCTGAACTACACTCATATCATCAAGTCTCACTGCTTCAGTGAAG 362
Db 570 CAGTAGAAGACTATATGAATCTCACACATACCATTCCAGTCTCACTGCTTCAGTGAAG 511
QY 363 GACAAAAGTGGCCAG-----CCTGGGGATGTTGCCAGCTTCTTGGAAAGTCAATTTGGTTC 416
Db 510 GACTATGGTCTCAGAAAAAATGTGGGATGCTGCCAAATCACTGGAAGTCAATTTGGCTC 451
QY 417 CAGTTGCTACTTCATTTCCAGTGAAG-AGAAGTTTGGTCTTAAGAGTGAAGCAACTGTG 475
Db 450 CAGCTGCTACTTCATTTCTACCAAGGAAGAACTTCTGGAGCACCAGTGAAGCAACTGTG 391
QY 476 TTCAGATGGGAGCACATTTGGTGTGTTTCAACACAGAACGACAGAGAAATTTTCATTGTCC 535
Db 390 TTCATATGGGGGCTCATCTGGTGGTGATCAATCACTGAACCGGAGCAAGATTTTCATCACCC 331
```

```
QY 536 AGCAGCTGAATGAGTGCATTTTCTTATTTTCTGGGCTTTTCAGACCCACAAAGCTAATA 595
Db 330 AGCAGCTGAATGAGTGCATTTTNTTACTTCTCGGTCTTTCGGATCCACAAGGTAATGGCA 271
QY 596 ATTGGCAATGGATTGATAAGACACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTG 655
Db 270 AATGGCAATGGATTGATGATGATCTCTTTCAGTCAAAATGTCAGGTCTGGCACCCCATG 211
QY 656 AGCCCAATCATTTCTGCAGAGCAATGTCTTCAATAGTCTTCTTGGAAACCTACAGGATGG 715
Db 210 AACCAATCTTCCAGAAGAGCGGTGTGTTTCAATAGTCTTACTTGAATCCTTCCAAATGG 151
QY 716 GCTGGAAATGATGTTATCTGTGAAACTAGAGGAATTCATATGCTGAGATCAATAAGATTT 775
Db 150 GCTGGAAATGATGTTTCTGTGTAGTAGTAAACACAATTCATATGTTGAAATGAAGAAGATTT 91
QY 776 ACCTATGAGT 785
Db 90 ACCTATGAGT 81

RESULT 3
US-09-833-381-995/c
; Sequence 995, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 995
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(528)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-995

Query Match 10.0%; Score 311.6; DB 10; Length 528;
Best Local Similarity 76.0%; Pred. No. 3.1e-49;
Matches 399; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

QY 183 AGAGAAAAGAGGCTGGTGTCTCCCTGAGACTCTGGTCTGTGGCTGGGATTTCCATTTGCAC 242
Db 525 AGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCTGAGCTGTGATTTTCCATTTTACT 466
QY 243 CCTCAGTGTCTTCATTTGCTGAGCTGTGTAGTAATTCATTTACATATGTTGAAC 302
Db 465 CTTGAGTACCTTGTTCATTTGCGAGCTGTGTGTGAGCTTACCAATTTATATGACACGCC 406
QY 303 TGSCAAAAGGCTGTCTGAACTACACTCATATCATCAAGTCTCACCTGCTTCAGTGAAG 362
Db 405 CAGTAGAAGACTATATGAATCTCACACATACCATTCCAGTCTCACCTGCTTCAGTGAAG 346
QY 363 GACAAAAGTGGCCAG-----CCTGGGGATGTTGCCAGCTTCTTGGAAAGTCAATTTGGTTC 416
Db 345 GACTATGGTCTCAGAAAAAATGTGGGATGCTGCCAAATCACTTGAAGTCAATTTGGCTC 286
QY 417 CAGTTGCTACTTCATTTCCAGTGAAGAGAGTGTGTTCTTAAGAGTGAAGCAACTGTG 476
Db 285 CAGCTGCTACTTCATTTCTACCAAGGAGAACTTCTGGAGCACCAGTGAAGCAACTGTG 226
QY 477 TGAGATGGGAGCACATTTGGTGTGTTTCAACACAGAACGACAGAGAAATTTTCATTGTCCA 536
Db 225 TCAGATGGGGCTCATCTGGTGGTGATCAATCACTACTGAAGCGGAGCAGAAATTTTCATCACCCA 166
```



RESULT 6

US-09-764-870-199

; Sequence 199, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT214

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 199

; LENGTH: 1091

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-870-199

Query Match 6.3%; Score 197; DB 10; Length 1091;

Best Local Similarity 67.8%; Pred.No. 7.1e-28;

Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

Qy 368 AGCTGCCAGCTGGGAGTCTGCCAGCTCTCTCGAAGTCAATTTGGTTCCAGTTGCTACT 427

Db 314 AAGAGACAGCTGGAGCTGTTGCCAAGAATTTGGAAGTCATTTAGTTCCAACTGCTACT 373

Qy 428 TCATTTCCAGTGAAGAGAGGTTTGGTCTAAGAGTGAGCAACTGTGTGAGATGGAG 487

Db 374 TTATTTCTACTCAATCAGCATCTTGGCAAGACACTGAGAAGGACTGTGCTAGAATGGAG 433

Qy 488 CACATTTGGTTGTTTCAACACAGACAGACAGAGCAATTTTCATTTGCCAGCAGCTGAATG 547

Db 434 CTCACCTGCTGTGATAAACACTCAAGAAGACAGCAGGATTTTCATCTCCAGAACTTCAAG 493

Qy 548 AGTCATTTCTTATTTCTGGGGCTTTTCAGACCCACACAGGTAAATAATTTGGCAATGCA 607

Db 494 AAGAATCTGCTTATTTTGGGGCTCTCCAGATCCAGAGGTGACGCACATTTGCCAATGGG 553

Qy 608 TTGATAAGACACCTTATAGAAAAATGTTCAGATTTTGGCACCTTAGGTGAGCCCAATCAT 667

Db 554 TTGATCAGACACCATACAAATGAAGTTCCACATTTGGCATCCAGTGAGCCCAAGTGC 613

Qy 668 CTCGACAGCAATGTGCTTCATATAGTCTTCTGGA---ACCTACAGGATGGGCTGGAATG 724

Db 614 CCAATCAGCGCTCGCTGTGTCTTAAATTTTCGTAATCACCACAAAGATGGGCTGGAATG 673

Qy 725 ATGTTATCTGTGAAACTAGAAAGAAATTCATATGTGAGATGAATTAAGATTACTATGAG 784

Db 674 ATGTTAATGTCTTGTGCTCTCAAGAGTCAAGTTGTGAGATGATCAAGTCCACTTATGAA 733

Qy 785 TAGAAGCTT 793

Db 734 CTGAACATT 742

RESULT 7

US-09-764-870-38

; Sequence 38, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT214

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 38

; LENGTH: 1096

; TYPE: DNA



FILE REFERENCE: SF0695B  
CURRENT APPLICATION NUMBER: US/09/862,802  
CURRENT FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US/09/111,470  
PRIOR FILING DATE: 1998-07-08  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 1418  
TYPE: DNA  
ORGANISM: unknown  
FEATURE:  
OTHER INFORMATION: mammalian nucleic acid  
NAME/KEY: CDS  
LOCATION: 279...992  
OTHER INFORMATION: protein coding sequence  
NAME/KEY: misc\_feature  
LOCATION: 1348  
OTHER INFORMATION: poly A addition motif  
US-09-862-802-7

Query Match 4.4%; Score 137.4; DB 9; Length 1418;  
Best Local Similarity 61.6%; Pred. No. 8.9e-17;  
Matches 257; Conservative 0; Mismatches 151; Indels 9; Gaps 2;  
QY 375 AGCCTGGGAGTGTGCCAGCTTCTTGGAAAGTCATTGGTTCGCTCCAGTTGCTACTTCAATTC 434  
Db 584 AGTCGGAGCTGTGCCCAAGAGATTGGAGGCTATTGGTTCCTCCACTGCTACTTGGTTCC 643  
QY 435 CAGTGA-----AGAGAAGTTTGGCTTAAGAGTGAGCAGAACTGTGTGAGATGGGAGC 488  
Db 644 CACAGTTTCTTCATCAGCATCTTGGAAACAGAGTGAGGAGAACTGCTCCCGCATGGGTGC 703  
QY 489 ACATTGGTGTGTCAACACAGAGAGCAGCAGAAATTCATTGTCACAGAGCTGAATGA 548  
Db 704 TCATCTAGTGGTGATCCAAAGCCAGAGAGCAGGATTTTCATCACTGGATCTTGGACAC 763  
QY 549 GTCAATTTCTTATTTCTGGGGCTTTCAGACCCACAAGGTAAATAATTTGGCAATGAT 608  
Db 764 TCATGCTGCTTATTTATAGGTTGTGGGATAC---AGGCCATCGGCAATGGCAATGGGT 820  
QY 609 TGATAGACACCTTATGAGAAAATGTACAGATTTTGGACCTAGTGTAGGCCCAATCATTC 668  
Db 821 TGATCAGACACCATATGAAGAAGTATACATTTCTGGCACAATGTGTAGAGCCAGAGTGG 880  
QY 669 TGCAGAGCAATGTGCTTCAATAGTCTTCTGGAACCTCAGAGATGGGCTGGAATGATGT 728  
Db 881 CAATGAAAATGTGCTACAAATATTTACCGTTGGAGACTGGATGGGCTGGAACGATAT 940  
QY 729 TATCTGTAAACTAGAGAAATTCATATGTGAGATGAATAAGATTTACCTATGAT 785  
Db 941 CTCTTGCAGTCTTAAACAGAAGTCAGTTTGTGATGAGAGAAAATAAACTTATGAAT 997

RESULT 12  
US-09-833-381-1339  
Sequence 1339, Application US/09833381  
Patent No. US20020132090A1  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
FILE REFERENCE: 5800-119  
CURRENT APPLICATION NUMBER: US/09/833,381  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 09/516,448  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 2050  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1339  
LENGTH: 758  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-833-381-1339

Query Match 3.9%; Score 120.6; DB 10; Length 758;  
Best Local Similarity 52.5%; Pred. No. 9.8e-14;  
Matches 321; Conservative 0; Mismatches 274; Indels 16; Gaps 2;  
QY 170 AACCTCAAAAGTACAGAGAAAGAGGC-TGGTTGCTCCCTCAGACTCTGGTCTGTGGCTGGG 228  
Db 18 AACCTCAAAAGTAAACTTGGAAAGGAGCATCATCCCCAGCTGCTCTTCGGTTATTGCTGTA 77  
QY 229 ATTTCCATTGCACTCTCAGTGTCTTCATTGTTCAGCTGTGTAGTAACCTTACCATTAT 288  
Db 78 GTTTTCATCTTACTTCTCAGTGTCTTTATTGCAAGTTGTTGTGACTCATCACAC 137  
QY 289 ACATATGGTGAACACTGGCAAAAGGCTGTCTGAACTACACTCATATATCATCAAGTCTCAC 348  
Db 138 TTTTCACGCTGTAAGAGAGAGCAGAGTGCACAAGTTAGAGCACCATTGCAAGGCTCAAA 197  
QY 349 TGTCTCAGTG-----AAGGACAAAGAGTGCACGCTGGGATGTTGCCCA 393  
Db 198 TGCATCAAGAGAGAAATCAGAACTTAAATCACTTAAGGGAGCACCTGGAACTGTTGTCT 257  
QY 394 GCTTCTTGGAAAGTCATTGGTTCAGTTCCTACTTCACTTCCAGTGAAGAGAGGTTTGG 453  
Db 258 ATTGACTGGAGAGCCTTCCAGTCCAACTGCTATTTTCCCTTACTGACAACAAGAGCTGG 317  
QY 454 TCTAAGAGTGTGAGAGCACTGTGTGAGATGGGAGCAGCAATTTGGTTGTTCACACAGAA 513  
Db 318 GCTGAGAGTGAAGGAACTGTTTCAGGGATGGGGGCCCATCTGATGACCATCAGCAGCGAA 377  
QY 514 GCAGAGCAGAAATTCATTGTCAGAGAGCTGAATGAGTCAATTTTCTTATTTTCGGGGCT 573  
Db 378 GCTGAGAGCAACTTATTTATTCAGTTTCTGGATAGAGCGGCTTCCCTATTTCTTGGACT 437  
QY 574 TCAGACCCACAAGGTAAATAATTTGCAATGGATTGATAAGACACCTTATGAGAAAAAT 633  
Db 438 AGAGATCAGAAATGCCAAAGCTCAGTGGCGTTGGTGACCAAGCCATTTAACCCACGC 497  
QY 634 GTCAGATTTTGGCACCTTAGTGTAGGCCCAATCACTCTGCAGAGCAATGTGCTTCAATAGTC 693  
Db 498 AGAGTATTTCTGGCATAAGAATGAACCGCACTCTCAGGGAGAAAACTGTGTGTCTT 557  
QY 694 TTTCTGAAACCTACAGAGATGGGCTGGAATGATGATCTATCTGTGAAACTAGAGAAATCA 753  
Db 558 GTTTATACCAAGATAATGGGCTGGAATGATGTTCCCTTGTAAAGCAAGTAGG 617  
QY 754 ATATGTGAGAT 764  
Db 618 ATTTGTAATAAT 628

RESULT 13  
US-09-965-529-40  
Sequence 40, Application US/09965529  
Publication No. US20020182671A1  
GENERAL INFORMATION:  
APPLICANT: LAL, Preeti  
APPLICANT: YUE, Henry  
APPLICANT: TANG, Y. Tom  
APPLICANT: BANDMAN, Olga  
APPLICANT: BURFORD, Neil  
APPLICANT: AZIMZAI, Yaida  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: LU, Dzung Aina M.  
APPLICANT: PATTERSON, Chandra  
TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0731 USA  
CURRENT APPLICATION NUMBER: US/09/965,529  
CURRENT FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315  
PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PERL Program  
SEQ ID NO 40

LENGTH: 968  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20020182671A1 1521513CB1  
US-09-965-529-40

Query Match 3.8%; Score 117.4; DB 9; Length 968;  
Best Local Similarity 54.4%; Pred. No. 4.1e-13;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

QY 170 AACCTCAAAAGTACAGAGAAAAGAGCGTGGTGTGCC---TGAGACTCTGTGTGGCTG 226  
DB 184 AAACACAATGCACAGAGAGAGGCTCTCTCTCCCAAATGTTCTTATGGACTGTTGCTG 243  
QY 227 GGATTTCCATTGCACTCCTCAGTCTGCTTCATTGTGAGCTGTGTAGTAACTTACCATT 286  
DB 244 GGATCCCCCATCCTATTCTCAGTGCCTGTTTCATCACCAAGATGTGTTGCACATTTCCGCA 303  
QY 287 TTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATATCATTCAAAGTCTCA 346  
DB 304 TCTTTCAACACCTGTCATGAGAAAAGTT---TCAGCTACCTGAGAATTTACAGAGCTCT 360  
QY 347 CCTGCTTCAGTGAAGGGACAAAAGTGGCCAGCTGGGATGTGCCCCAGCTTCTTGGAGT 406  
DB 361 CCTGCTACAATTATGGATCA---GGTTCAGTCAAGAATTGTTGTCCATTGAACTGGCAAT 417  
QY 407 CATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAGAGTTTGGTCTAAAGAGTACG 466  
DB 418 ATTTTCAATCCAGCTGCTACTTCTTTCTACTGACACCATTTCTCTGGCGGTTAAGTTTAA 477  
QY 467 AGAAGTGTGGTGGAGTGGACACATTTGGTGTGTTTCAACACAGAGACGACGAAAT 526  
DB 478 AGAAGTGTTCAGCATGGGGCTCACCTGGTGGTTATCAACTACAGAGAGACGAGAAAT 537  
QY 527 TCAATGTCACAGAGCTGAATGAGTCAATTTCTTTATTTCTGGGCTTTTCAGACCCACAAG 586  
DB 538 TCCCTTCTCAACAAGAACTAAATGAGAGAGTTTATTTATGGACTGTCAGACCCAGGTTG 597  
QY 587 GTAATAAATGGAATGGCAATGGATTGATAGACACCTTATGAGAAAATGTCAGATTTGGC 646  
DB 598 TCGAGGGTCAGTGGCAATGGGTGGACGGCACACCTTTGACAAAAGTCTCTGAGCTTCTGGG 657  
QY 647 ACCTAGGTGACCCCAATCA-----TTCGCGACAGCAATGTGCTCAATAGTCTCTCGGA 700  
DB 658 ATGTAGGGGACCCCAACACATAGCTACCTCCCTGGAGGACTGCGCCACCATGAGAGACTCTT 717  
QY 701 AACCTACAGGATGGGGCTGGGAATGATGTTATCTGTGAAACTAGAGGAATTCATATGTG 760  
DB 718 CAACCCCAAGCAAAATTTGGAATGATGTAACCTGTTTCTCCTCAATTTATTTCCGGATTTGTG 777  
QY 761 AGATG 765  
DB 778 AATG 782

RESULT 14  
US-09-905-291A-376  
; Sequence 376, Application US/09905291A  
; Patent No. US20020160374A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerriksen, Mary E.

APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/905,291A  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 376  
LENGTH: 997  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-905-291A-376

Query Match 3.8%; Score 117.4; DB 9; Length 997;  
Best Local Similarity 54.4%; Pred. No. 4.2e-13;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

QY 170 AACCTCAAAAGTACAGAGAAAAGAGCGTGGTGTGCC---TGAGACTCTGTGTGGCTG 226  
DB 128 AAACACAATGCACAGAGAGAGGATGCTCTCTTCCCAAATGTTCTTATGGACTGTTGCTG 187  
QY 227 GGATTTCCATTGCACTCCTCAGTCTGCTTTCATTGTGAGCTGTGTAGTAACTTACCATT 286  
DB 188 GGATCCCCCATCCTATTCTCAGTGCCTGTTTCATCACCAAGATGTGTTGTGACATTTCCGA 247  
QY 287 TTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATATCATTCAAAGTCTCA 346









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:47:16 : Search time 2991.64 Seconds  
(without alignments)  
16857.906 Million cell updates/sec

Title: US-09-766-511B-51  
Perfect score: 3114  
Sequence: 1 cttaatgttggagtcctctt.....tttaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	366.8	11.8	383	17	AQ091895 HS_3014.B
2	366.8	11.8	1900	11	BC001229 Homo sapi
3	366.2	11.8	563	10	AW576568 UI-HF-BN0
4	365.8	11.7	582	13	BM505485 ih12e08.x
5	364.6	11.7	732	14	BM972330 UI-Cf-EC1
6	364.6	11.7	764	14	BQ009055 UI-H-E10-

7	364.2	11.7	511	9	A1745635
8	363	11.7	1030	13	BM473183
c 9	361.4	11.6	588	13	BI868214
c 10	360.6	11.6	1079	13	BM463804
11	359.8	11.6	538	10	AW576183
12	358.2	11.5	548	9	A1684386
13	358.2	11.5	577	14	BM968219 UI-CF-DU1
14	355.6	11.4	725	14	BM715473 UI-E-EJ0-
15	355	11.4	496	9	AA557129 n17Aa09.s
16	354	11.4	528	13	BM661657 UI-E-CK1-
17	354	11.4	735	14	BM678541 UI-E-E00-
18	354	11.4	736	14	BM679660 UI-E-E00-
19	354	11.4	762	14	BM679029 UI-E-E00-
20	352.2	11.3	519	14	C75436
21	350.2	11.2	496	9	A1025096
22	350.2	11.2	510	9	A1018792
23	349.8	11.2	596	9	AA172069
24	349.8	11.2	599	9	AA172043
25	349.6	11.2	488	10	BE350097
26	349.2	11.2	506	10	AW148744
27	349	11.2	487	10	BE302211
28	347.2	11.1	578	9	AU149516
29	346.8	11.1	600	9	AU145485
30	346.4	11.1	652	9	AA172077
c 31	344.6	11.1	534	14	BM754950
32	344.2	11.1	590	9	AU157092
33	343.6	11.0	504	12	BF603486
34	343.6	11.0	543	9	AU152484
35	342.6	11.0	477	9	A1924217
36	341.8	11.0	477	9	A1422372
c 37	341.6	11.0	480	14	BM832889
38	340.6	10.9	545	10	BE177944
39	338.8	10.9	553	13	BM145299
40	338.8	10.9	573	9	AU151353
41	338.6	10.9	477	9	A1040207
42	338.6	10.9	477	9	A1660533
43	338	10.9	545	13	BG942747
44	337.8	10.8	535	9	AU153279
45	335.4	10.8	465	9	A1246452

ALIGNMENTS

RESULT 1	AQ091895	HS_3014_B2_C05_T7	CIT	Approved Human Genomic Library D Homo sapiens genomic clone	3014 Col=10 Row=F, DNA sequence.
AQ091895	AQ091895	AQ091895.1	GI:3462982		
LOCUS	HS_3014_B2_C05_T7	CIT	Approved Human Genomic Library D Homo sapiens genomic clone	3014 Col=10 Row=F, DNA sequence.	
DEFINITION	AQ091895	AQ091895.1	GI:3462982		
ACCESSION	AQ091895	AQ091895.1	GI:3462982		
VERSION	AQ091895	AQ091895.1	GI:3462982		
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 383)				
	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.				
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				
MEDLINE	99380589				
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 3014 row: F column: 10				

Class: BAC ends  
High quality sequence stop: 383.  
Location/Qualifiers  
1. .383  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-3014 Col-10 Row-F"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"  
BASE COUNT 110 a 72 c 49 g 148 t 4 others  
ORIGIN  
Query Match 11.8%; Score 366.8; DB 17; Length 383;  
Best Local Similarity 97.1%; Pred. No. 1.4e-46;  
Matches 371; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 843 TTAATAATGAACCCATCATGAATGATAATTTCTTCTGATTTACACATAATCCCTTAT 902  
Db 1 TTAANATTGAACCCATCATGAATGATAATTTCTTCTGATTTACACATAATCCCTTAT 60  
QY 903 GTTATAGAGTTTCACAGAAATGAAGATACCTGTTTCCCTTTAATCAATCTTCTCGTTT 962  
Db 61 GTTATAGAGTTTCACAGAAATGAAGATACCTGTTTCCCTTTAATCAATCTTCTCGTTT 120  
QY 963 CCTCTTTTCCATTAATGATAGAAATGCACCCCTTCTCTCTTTGTTCCATTTCTTCACTGT 1022  
Db 121 CCTCTTTTCCATTAATGATAGAAATGCACCCCTTCTCTCTTTGTTCCATTTCTTCACTGT 180  
QY 1023 TATTCATTTTTTCTTCTTCCACTTCATTTACACAAATATTTATGTTTTCAGAGACGT 1082  
Db 181 TATTCATTTTTTCTTCTTCCACTTCATTTACACAAATATTTATGTTTTCAGAGACGT 240  
QY 1083 ACTATTTGTTTGTAGAGATTTATAAGGAGTATCTTTTCARAAATTTATGACITTCCTT 1142  
Db 241 NACTATTTTGTGTAGAGATTTATAAGGAGTATCTTTTCARAAATTTATGACITTCCTT 300  
QY 1143 CCTCAATATACCATAAAGAAATCTTTTGGTCAAGATGGTAGTTGGAAGTACAATCATCT 1202  
Db 301 CCTCAATATACCATAAAGAAATCTTTTGGTCAAGATGGTAGTTGGAAGTACAATCATCT 360  
QY 1203 GAAGGCTGACAAGAGTTGAAA 1224  
Db 361 GAAGGCTGACAAGAGTTGAAA 382  
RESULT 2  
BC001229/c 1900 bp mRNA linear HTC 31-JAN-2002  
LOCUS Homo sapiens, Similar to RAD21 (S. pombe) homolog, clone  
DEFINITION IMAGE:3455406, mRNA.  
ACCESSION BC001229  
VERSION BC001229.1 GI:12654772  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1900)  
Strausberg,R.  
Direct Submission  
Submitted (11-DEC-2000) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: http://www-shgc.stanford.edu  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAC Plate: 4 Row: 1 Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 7705806  
This clone has the following problem: no 5' EST match.  
FEATURES  
Location/Qualifiers  
source 1. .1900  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3455406"  
/tissue\_type="Cervix, carcinoma"  
/clone\_lib="NIH\_MGC\_12"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 632 a 294 c 362 g 612 t  
ORIGIN  
Query Match 11.8%; Score 366.8; DB 11; Length 1900;  
Best Local Similarity 86.4%; Pred. No. 5.8e-47;  
Matches 472; Conservative 0; Mismatches 47; Indels 27; Gaps 5;  
QY 2477 TTTCATTTCTCTGCTCCAAAATATTTTTCACAGCTTATAACACAACTTTTATTAGA 2536  
Db 1896 TTTCATTTCTCTGCTCCAAAATATTTTTCACAGCTTATAACACAACTTTTATTAGA 1837  
QY 2537 AAAGTTATACATAACAGCATCAACTATTTTCAAG-----AACCCAATAAGCAA 2586  
Db 1836 AAAGTTATACATAACAGCATCAACTATTTTCAAGACAATATTAAACCGATAGCAA 1777  
QY 2587 CAAAAACAGACTAACAAAATGTGTACAGAACTTAAGACCTTTCTTAAATCAAAAT 2646  
Db 1776 CAAAAACAGACTAACAAAATGTGTACAGAACTTAAGACCTTTCTTAAATCAAAAT 1717  
QY 2647 TCAATTATCTACAAATGTCATTTTACAAACAGGAAACCTCCATGTTTACAGGCATGCA 2706  
Db 1716 TCAATTATCTACAAATGTCATTTTACAAACAGGAAACCTCCATGTTTACAGGCATGCA 1657  
QY 2707 TATTGAAATTAAGCTGCAATAGC-TTTTATATACAAATTCGCTCTCAAGAAATGAATC 2765  
Db 1656 TATTGAAATTAAGCTGCAATAGCATTATATACAAATACCACCTGGAAGAACTGAATC 1597  
QY 2766 ATTAAGACAGTAATTAGGAGTTCACAAAATTTAAACATTTTCAGTAAATTTTAAATTTATG 2825  
Db 1596 ATTAAGACAGTAATTAGGAGTTCACAAAATTTAAACATTTTCAGTAAATTTTAAATTTATG 1537  
QY 2826 TCCTCAATAATTTTAAATTTTCAAGTTCAGTTTCAAAAGTGA---TTTTTCCCAACA 2882  
Db 1536 GGT-----ATACACTGAAGTCTGAGTTTCAAAAGTGAATTTTTCCTCCCAACA 1489  
QY 2883 AGGTGCCAACACTTAAGCTAGAGCTTTTCAAGTGTAACTTTTGCCT-AAAAGTTTAAGACAT 2941  
Db 1488 AAGTTTCAACACTTAAGCTAGAGCTTTTCAAGTGTAACTTTTGCCTTAAAGTTTAAGACAT 1429  
QY 2942 ATTCTGAGAAATCAATATAGTCACATGATTTCTGATGCTATCTGCTCTGTATAACAAG 3001  
Db 1428 ATTCTGAGAAATCAATATAGTCACATGATTTCTGATGCTATCTGCTCTGTATAACAAG 1369  
QY 3002 ATTCTCA 3007  
Db 1368 TCCTTCA 1363  
RESULT 3  
AW576568  
LOCUS AW576568 563 bp mRNA linear EST 15-MAR-2000  
DEFINITION UI-HF-BN0-alu-a-09-0-UI-sl NIH\_MGC\_50 Homo sapiens .cdna clone

Db	356	 AAGTCGAGTTTCAAAAGTGATTTTTTTCCACAAAAGTTTCAACACACTTAAGCTAGAA	415
QY	2906	CTTTTCAGTGTAACTTTGGCCT-AAAAAGTTAAAGACATATTCTTGAGAATCATATAATAGTCAC	2964
Db	416	CTTTTCAGTGTAACTTTGGCCTAAAAAGTTAAAGACATATTCTGTATATCATACAAGTCAC	475
QY	2965	ATGATTTCTGATCGTATPCTGCTCTGTTTAATAACAAAGATTCA	3007
Db	476	ATGATTTCTGATCGTATPCTGCTGTTTAATAAAAGCTTTTA	518
RESULT 4			
BM505485			
LOCUS		582 bp mRNA linear	EST 14-FEB-2002
DEFINITION		ihl2e08.x1 Human insulinoma Homo sapiens cDNA 3', mRNA sequence.	
ACCESSION		BM505485	
VERSION		BM505485.1 GI:18668752	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 582)	
		Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,	
		Lemishka, I., Scearcie, M., Brestelli, J., Gradwohl, G., Clifton, S.,	
		Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blisstein, A.,	
		Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas	
		, M., Gibbons, M., McGann, R., Cole, R., Tsagareishvili, R., Williams, T.	
		, Jackson, X., and Bowers, Y.	
TITLE		Endocrine Pancreas Consortium	
JOURNAL		Unpublished (2000)	
COMMENT		Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue	
		Endocrine Pancreas Consortium	
		Harvard University, Howard Hughes Medical Institute	
		Dept Of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,	
		MA 02138	
		Tel: 617-495-1812	
		Fax: 617-495-8557	
		Email: dmelton@biocph.harvard.edu	
		Library was constructed by Dr. J. Ferrer In vivo mass-excised to	
		pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington	
		University Genome Sequencing Center For information on obtaining a	
		clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)	
		Seq primer: -400P from Gibco	
		High quality sequence stop: 473.	
FEATURES		Location/Qualifiers	
source		1..582	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone_lib="Human insulinoma"	
		/tissue_type="insulinoma"	
		/lab_host="Dhl0B (phage-resistant)"	
		/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:	
		xhoI; Site_2: EcoRI; Constructed with lambda ZAPII system	
		(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to	
		pBluescript SK- by Dr. H. Inoue following the Washington	
		University protocol	
		(http://genome.wustl.edu/est/lambda_protocol.shtml).	
		Please contact Hiroshi Inoue, MD/PhD for further	
		information on this library (Metabolism Division, Permutt	
		Laboratory, Washington University School of Medicine, Box	
		8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this	
		is a Washington University pancreas EST project library."	
BASE COUNT	221 a	95 c 72 g 194 t	
ORIGIN			
Query Match		11.7%; Score 365.8; DB 13; Length 582;	
Best Local Similarity		88.7%; Pred.No. 1.6e-46;	
Matches 462; Conservative		0; Mismatches 32; Indels 27; Gaps 5;	
QY 2502	TTTTTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATAACACAGCATCAA	2561	



```

VERSION      B009055.1  GI:19733956
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 764)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: Dr. Jose Mercuende
             cDNA Library Arranged by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Seq primer: M3 FORWARD
             POLYA=yes.

FEATURES     Location/Qualifiers
             1..764
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image="5840869"
             /clone_lib="NCI-CGAP_E10"
             /tissue_type="Chondrosarcoma"
             /dev_stage="Adult"
             /lab_host="DH10B (Life Technologies)"
             /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
             with a modified polylinker; Site.1: EcoR I; Site.2: Not I;
             NCI-CGAP_E10 is a cDNA library containing the following
             tissue(s): Chondrosarcoma. The library was constructed
             according to Bonaldo, Lennon and Soares, Genome Research,
             6:791-806, 1996. First strand cDNA synthesis was primed
             with an oligo-dT primer containing a Not I site. Double
             stranded cDNA was ligated to an EcoR I adaptor, digested
             with Not I, and cloned directionally into pT7T3-Pac
             vector. The oligonucleotide used to prime the synthesis of
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dT)18 tail. The
             sequence tag for this library is ACATTGTCAC.
             TAG_LTB=UI-R-E10
             TAG_TISSUE=chondrosarcoma
             TAG_SEQ=ACATTGTCAC"

BASE COUNT   272 a 131 c 107 g 253 t 1 others
ORIGIN
Query Match 11.7%; Score 364.6; DB 14; Length 764;
Best Local Similarity 88.3%; Pred. No. 2e-46;
Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;

Qy 2500 TATTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATC 2559
Db 1 TTTTATTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACATACATC 68
Qy 2560 AACTATTTTTCAAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGT 2609
Db 69 AACTATTTTTCAAGAACAAATATTAACCCGATAGACAAAAACCCAGACTAACAAAATGT 128
Qy 2610 GTAACAGAAACTAATGACCTTCTTAAATCAACATTCATATCTACATATCTATTT 2669
Db 129 GTACAGAAACTAATGACCTTCTTAAATCAACATTCATATCTACATATCTATTT 188
Qy 2670 ACAACAGGGGAAACTCCATGGTTTACAGGCATGTCATATTGAAATTAAGCTGCAATAG 2729
Db 189 ACAACAGGGGAAACTCCCTGGTTTACAGGCATCATATATTGAATATAAAGCTGCAATAG 248
Qy 2730 C-TTTTATACAAATATCGCTCTCAAGAAATGAATCATTTAAGACAGATAATTAGAGTTC 2788
Db 249 CAATTTTATACAAATATGACACTCTGAAGAACTGAATCATTTAAGACAGATAATTAGAGTTC 308

Qy 2789 ACAAAATTTAAACATTTTCACGTAATTTTAAATATTCTTCTTCAATAATTTTAAATTTT 2848
Db 309 ACAAAATTTAAACATTTTCACATAATTTTAAATATTGGGT-----ATACACTG 356
Qy 2849 AAGTCTGAGTTTCAAAAAGTGA---TTTTTCCCAACAAGGTGCCAACACTTAAGCTAGAG 2905
Db 357 AAGTCTGAGTTTCAAAAAGTGAATTTTTCCTCCCAAAAAGTTTCAACACTTAAGCTAGAA 416
Qy 2906 CTTTCAGTGTTAACCTTTGCCCCCT-AAAAGTTAAGACATATTTCTGAGATCATATAATAGTCAC 2964
Db 417 CTTTCAGTGTTAACCTTTGCCCCCTAAAAGTTAAGACATATTTCTGATATATCATACAGTCAC 476
Qy 2965 ATGATTTCTGATGCTATCTGCTCTGTTTAAATAACAAAGATTTC 3007
Db 477 ATGATTTCTGATGCTATCTGCTCTGTTTAAATAAAGTCTTTA 519

RESULT 7
LOCUS      AI745635
DEFINITION tr24a09.xl NCI-CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219224 3',
             mRNA sequence.
ACCESSION  AI745635
VERSION    AI745635.1 GI:5113923
KEYWORDS  EST.
SOURCE     human.
ORGANISM  Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 511)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone Distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Seq primer: -400P from Gibco
             High quality sequence stop: 414.
             Location/Qualifiers
             1..511
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image="2219224"
             /clone_lib="NCI-CGAP_Ov23"
             /tissue_type="tumor, 5 pooled (see description)"
             /lab_host="DH10B"
             /note="Organ: ovary; Vector: pCMV-SPORT6; Site.1: SalI;
             Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
             Average insert size 1.35 kb. Tumor types include: mixed
             Mullerian tumor, papillary serous, clear cell, spindle
             cell. All are primary tumors, metastasis positive. Life
             Technologies catalog #: 11534-013"

BASE COUNT   199 a 87 c 57 g 168 t
ORIGIN
Query Match 11.7%; Score 364.2; DB 9; Length 511;
Best Local Similarity 88.5%; Pred. No. 2.9e-46;
Matches 461; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

Qy 2502 TTTTATTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATACACAGCATCAA 2561
Db 1 TTTTATTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATACATACATCAA 60
Qy 2562 CTATTTTTCAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGTCT 2611
Db 1 TTTTATTTTTCAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGTCT 2611

```

Db	61	CTATTTTCAAGAACAAATATTAAACCCGATAGCAACAAAAACCAGACTAACAAAAATGTGT	120
Qy	2612	AACAAGAACTAATGACCTTTCTTAAATCAAACATTCATTAATCTACAATGTCTATTAC	2671
Db	121	AACAAGAACTAATGACCTTTCTTAAATCAAACATTCATTAATCTACAATGTCTATTAC	180
Qy	2672	AAACAGGAAATCCATGGTTTACAGGCATGTCATATTGAAATAAAGCTGCAATAGC-	2730
Db	181	AAACGGGAAACTCCTTGGTTTACAGGCACATCATATTGAATATAAAGCTGCAATAGCA	240
Qy	2731	TTTTTATACAAATTCGCTCTCAAGAAAAATGAATCATTAAGACAGTAATTAAGAGTTCAC	2790
Db	241	ATTTTATACAAATTAACACTCTGAAGAATCAATCATTAACAGTAATTAACAGTTCAC	300
Qy	2791	AAATTTAAACATTCACGTAATTTTAAATTAATTTGTTTCAATAATTTTAAATTAATTGAA	3850
Db	301	AAATTTAAACATTCACATAATTTTAAATTAATTTGGGT-----ATACACTGAA	348
Qy	2851	GTCTGAGTTTCAAAAGTGA--TTTTCCTCCACAAGGTCGCAACACTTAAGCTAGAGCT	2907
Db	349	GTCTGAGTTTCAAAAGTGAATTTTTCCTCCACAAGGTTTCAACACTTAAGCTAGAGCT	408
Qy	2908	TTCACTGTTAACTTTGCGCT-AAAAGTTAAGACATATTTCTGAGAATCATAATAGTCACAT	2966
Db	409	TTCACTGTTAACTTTGCGCTAAAAGTTAAGACATATTTCTGAGAATCATAATAGTCACAT	468
Qy	2967	GATTTCTGATGCTATCGCTCTGTTAATAACAAGATTTCAC	3007
Db	469	GATTTCTGATGCTATCGCTCTGTTAATAATAAAGTCTTTA	509
RESULT 8			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			

Best Local Similarity 88.1%; Pred. No. 3.le-46;			
Matches 461; Conservative 0; Mismatches 35; Indels 27; Gaps 5;			
Qy	2500	TATTTTTCACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATC	2559
Db	9	TTTTTTTTTACCGTTTAAACAACTTTTATTAGAAAAGTTATACATACATAGCATC	68
Qy	2560	AACTATTTTCAAG-----AACCCAATAAGCAACAAAAACCAGACTAACAAAAATGT	2609
Db	69	AACTATTTCAAGAACAAATATTAAACCCGATAGCAACAAAAACCAGACTAACAAAAATGT	128
Qy	2610	GTAACAAGAACTAATGACCTTTCTTAAATCAAACATTCATTAATCTACAATGTCTATT	2669
Db	129	GTAACAAGAACTAATGACCTTTCTTAAATCAAACATTCATTAATCTACAATGTCTATT	188
Qy	2670	ACAAACAGGAAATCCATGGTTTACAGGCATGTCATATTGAAAATAAAGCTGCAATAG	2729
Db	189	ACAAACGGGAAACTCCTTGGTTTACAGGCACATCATATTGAATATAAAGCTGCAATAG	248
Qy	2730	C-TTTTTATACAAATATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGGAGTTC	2788
Db	249	CAATTTTATACAAATTAACACTCTGAAGAACTGAATCATTAACACAGTAATTAGGAGTTC	308
Qy	2789	ACAAATTTAAACATTTCACTTAATTTTAAATTAATTTGCTTCAATAATTTTAAATTTATG	2848
Db	309	ACAAATTTAAACATTTCACTTAATTTTAAATTAATTTGGGT-----ATACACTG	356
Qy	2849	AAGCTGAGTTTCAAAAGTGA--TTTTCCTCCACAAGGTCGCAACACTTAAGCTAGAG	2905
Db	357	AAGCTGAGTTTCAAAAGTGAATTTTTCCTCCACAAGGTTTCAACACTTAAGCTAGAA	416
Qy	2906	CTTTCAGTTTAACTTTGCCCT-AAAAGTTAAGACATATTTCTGAGAATCATTAATAGTTCAC	2964
Db	417	CTTTCAGTTTAACTTTGCCCTAAAAGTTAAGACATATTTCTGATATATACATACAGTTCAC	476
Qy	2965	ATGATTTCTGATGCTATCGCTCTGTTTAAATAACAAGATTTCAC	3007
Db	477	ATGATTTCTGATGCTATCGCTCTGTTTAAATAAAGTCTTTA	519
RESULT 9			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			





**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. Tissue procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M.B. Soares Lab  
cDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.

<b>FEATURES</b>					
source		Location/Qualifiers			
		1..538			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:3065073"			
		/clone_lib="NIH_MGC_50"			
		/tissue_type="lymph"			
		/cell_type="germinal center B cells"			
		/cell_line="MGC85"			
		/lab_host="DH10B (JTI)"			
		/note="vector: pT73-Pac; site_1: NotI; site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, ph.D. and M. Bento Soares, Ph.D."			
BASE COUNT		201 a	91 c	61 g	185 t
<b>ORIGIN</b>					
Query Match		11.6%; Score 359.8; DB 10; Length 538;			
Best Local Similarity		87.8%; Pred. No. 1.3e+45;			
Matches 459; Conservative		0;	Mismatches 37;	Indels 27;	Gaps 3;
QY	2500	TATTTTTTTTACAGCTTATAACACAACCTTTATTAGAAAAGTTATACATAAACACAGCATC	2559		
Db	8	TTTTTTTTTTACAGCTTATAACACAACCTTTATTAGAAAAGTTATACATAACATAGCATC	67		
QY	2560	AAC TATTTTTC AAG-----AACCCAATTAAGCAACAAAAACCAGACTAACA AAATGT	2609		
Db	68	AAC TATTTTTC AAGAACAATATTA AACCGGATTAAGCAACAAAAACCAGACTACAAATGT	127		
QY	2610	GTAACAGAAACTAATGACCTTCTAAATCAAA CATTC AATTTATCTACAATGCTCTATTT	2669		
Db	128	GTAACAGAAACTAATGACCTTCTAAATCAAA CATTC AATTTATCTACAATGCTCTTTT	187		
QY	2670	ACAAACAGGGAAACCTCCATGGTTTACAGGCTGTCATATTGAAAATAAGCTGCAATAG	2729		
Db	188	ACAAACGGGAAACCTCCTGGTTTACAGGCATCATATTGAATATTAAGCTGCCAATAG	247		
QY	2730	C-TTTTATATACAATTTATCGCTCTCAAGAAATGAATCATTTAAGACAGTAATTAGGAGTTC	2788		
Db	248	CAATTTTATACAAATTTACCACCTCTGAAGAACTGAATCATTTAAAACACAGTAATTCAGAGTTC	307		
QY	2789	ACAAATTTTAAAACATTTTCACCTGAATTTTAAATATTGCTCTCAATATTTTAAATTTATTG	2848		
Db	308	ACAAATTTTAAAACATTTTCACATATTTTAAATATTG GGT-----ATACACTG	355		
QY	2849	AAGCTCGAGTTTCAAAGTGA---TTTTTCCCACAAAGTGCCAACTTAAGCTAGAG	2905		
Db	356	AAGCTCGAGTTTCAAAGTGAATTTTTTTTCCCACAAAGTTTTTCAACACTTAAGCTAGAA	415		
QY	2906	CTTTCAGTGTTAACTTTCCCGCT-AAAAGTTAAGCACATATCTGAGATCATATAAGTCAC	2964		
Db	416	CTTTCAGTGTTAACTTTCCCGCTAAAAGTTAAGCACATATTTCTGATTAATCATTAACAGTCAC	475		
QY	2965	ATGATTTCTGATGCTATCTGCTCTGTTTAATAACAAAGATTTC A 3007			

Db 476 ATGATTTCTGATGCTATCTGCCTGTTAATAATAGACTCTTTA 518

RESULT 12

AI684386

LOCUS

tc96h02.x1 NCI\_CGAP\_CLL1 Homo sapiens cDNA clone IMAGE:2074035 3',

DEFINITION

mRNA sequence.

ACCESSION

AI684386

VERSION

AI684386.1

KEYWORDS

GI:4895680

SOURCE

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 548)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL

Tumor Gene Index

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapb-r@mail.nih.gov](mailto:cgapb-r@mail.nih.gov)

Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

cDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Insert length: 707 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 471.

Location/Qualifiers

1..548

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2074035"

/clone\_lib="NCI\_CGAP\_CLL1"

/tissue\_type="B-cell, chronic lymphocytic leukemia"

/lab\_host="DH10B"

/note="Vector: pT7f3D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTCAAGTGGGCGGCATGCTGCTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7f3 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

BASE COUNT 214 a 91 c 64 g 179 t

[illegible]



Seq primer: M13 Reverse.

RESULT 15	AA557129	496 bp	mRNA	linear	EST 09-SEP-1997
LOCUS	AA557129				
DEFINITION	n174a09.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1056376 3', mRNA sequence.				
ACCESSION	AA557129				
VERSION	AA557129.1	GI:2327606			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 496)				
TITLE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-f@mail.nih.gov">cgapbs-f@mail.nih.gov</a> Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://www.bio.llnl.gov/bbrp/image/image.html">www.bio.llnl.gov/bbrp/image/image.html</a> Insert length: 656 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 474.				

```

1. .496
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:1056376"
/clone_lib="NCI_CGAP_Br2"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/note="Vector: pT7c3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7c3
vector. This library is the normalized version of
NCI_CGAP_Brl.1. Library was constructed by Bento Soares
and M. Fatima Bonaldo."
BASE COUNT      198 a      86 c      57 g      155 t
ORIGIN
Query Match      11.4%      Score 355;  DB 9;  Length 496;
Best Local Similarity 88.8%;  Pred. No. 7.5e-45;
Matches 450;  Conservative 0;  Mismatches 30;  Indels 27;  Gaps 5;

QY      2510  ACAGCTTTAACAACAACCTTTTATTAGAAAAAGTTATACATAACACAGACTCAACTATTTTC 2569
|||||
Ddb      1  ACAGCTTATAACAACACTTTTATTAGAAAAAGTTATACATAACATAGCATCAACTATTTTC 60
|||||

QY      2570  AAG-----RACCATAAGCACACAAACCAGACTTAACAAAATGTTTAACAAGAA 2619
|||||
Ddb      61  AAGAACAAATATTAAACCCGATAGCAACAAAACAGACTTAACAAAATGTTTAACAAGAA 120
|||||

```

Qy 2620 ACTAATGACCTTCTAAATCAAACTTCAATTATCTACAATGCTCTATTACAAACAGG 2679  
Db 121 ACTAATGACCTTCTAAATCAAACTTCAATTATCTACAATGCTCTATTACAAACAGG 180  
Qy 2680 AAAACTCCATGGTTTACAGGCGATGTCATATTGAAAATAAAGCTGCAATAGC-TTTTATA 2738  
Db 181 AAAACTCCCTGGTTTACAGGCGACATCATATTGAATATAAGCTGCAATAGCAATTTATA 240  
Qy 2739 CAATTATCGCTCTCAGAAATGAATCAATTAAGACAGTAATTAGGAGTTTCACAAATTTAA 2798  
Db 241 CAATTACCACTCTCAGAAACTGAATCAATTAAGACAGTAATTAGGAGTTTCACAAATTTAA 300  
Qy 2799 AACATTTTCAGCTAATTTTAAATTAATCTCTCAATTAATTTTAAATTAATGAAGCTGAGT 2858  
Db 301 AACATTTTCACATTAATTTTAAATTAATTTGGGT-----ATACACTGAAGCTGAGT 348  
Qy 2859 TTCAAAAGTGA---TTTTTCCACAAAGGTGCCAACACTTAAGCTAGAGCTTTTCAGTGT 2915  
Db 349 TTCAAAAGTGATTTTTCACAAAGGTGCCAACACTTAAGCTAGAGCTTTTCAGTGT 408  
Qy 2916 TAACCTTTGCCCT-AAAAGTTAAGACATATTTCTGAGAAATCATAATAGTCACATGATTTCTG 2974  
Db 409 TAACCTTTGCCCTAAAAGTTAAGACATATTTCTGAGAAATCATAATAGTCACATGATTTCTG 468  
Qy 2975 ATGCTATCTGCTCTGTTTAATAACAAAG 3001  
Db 469 ATGCTATCTGCTCTGTTTAATAACAAAG 495

Search completed: February 17, 2003, 11:47:10  
Job time : 3022.64 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:46:11 ; Search time 5370.63 Seconds  
(without alignments)  
16874.395 Million cell updates/sec

Title: US-09-766-511b-51  
Perfect score: 3114  
Sequence: 1 cttaatgttggaagtctctt.....tttaaaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pi:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_man:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2446.6	78.6	109320	9	AC092746	AC092746 Homo sapi
C 2	379.2	12.2	1227	6	AX155225	AX155225 Sequence
C 3	379.2	12.2	1227	10	AF240357	AF240357 Mus muscu
C 4	379.2	12.2	1254	10	BC023008	BC023008 Mus muscu
C 5	364.6	11.7	3708	6	AX014883	AX014883 Sequence
C 6	357.2	11.5	156441	9	AP004074	AP004074 Homo sapi
C 7	356.2	11.4	2209	9	AK097915	AK097915 Homo sapi
C 8	356.2	11.4	3647	9	HUMORF005	D38551 Human mRNA
C 9	346.6	11.1	193100	2	AC016729	AC016729 Homo sapi
C 10	346.6	11.1	193364	9	AC087350	AC087350 Homo sapi
C 11	317.6	10.2	37652	9	HS085A3	Z78021 Human DNA s
C 12	317.6	10.2	56583	9	HS347M6	295327 Human DNA s
C 13	283.8	9.1	1312	6	AX155223	AX155223 Sequence
C 14	283.8	9.1	1313	9	AF293615	AF293615 Homo sapi
C 15	280.6	9.0	851	9	AF325459	AF325459 Homo sapi
C 16	276.8	8.9	1125	10	AF240358	AF240358 Mus muscu
C 17	272.8	8.8	827	6	AX357481	AX357481 Sequence
C 18	269	8.6	800	6	AX357483	AX357483 Sequence
C 19	257.8	8.3	3298	6	E11759	E11759 Structural
C 20	257.8	8.3	3298	10	MUSB14GA	D37790 Mouse mRNA
C 21	255.4	8.2	684	6	AX098173	AX098173 Sequence
C 22	244	7.8	1125	10	AF240359	AF240359 Mus muscu
C 23	235	7.5	428	11	G23519	G23519 human STS W
C 24	233.2	7.5	739	9	AF325460	AF325460 Homo sapi
C 25	224.4	7.2	3929	6	AX305440	AX305440 Sequence
C 26	224.4	7.2	3929	10	MUSF9	D49429 Mouse NCBP-
C 27	201.2	6.5	444	6	AX456975	AX456975 Sequence
C 28	200.4	6.4	402	6	AX456962	AX456962 Sequence
C 29	197	6.3	1133	9	AF109146	AF109146 Homo sapi
C 30	197	6.3	1195	6	AX400064	AX400064 Sequence
C 31	197	6.3	1271	9	HS133532	AJ133532 Homo sapi
C 32	195.4	6.3	1305	9	AF067800	AF067800 Homo sapi
C 33	194.2	6.2	615	9	AF200738	AF200738 Homo sapi
C 34	186	6.0	1308	9	AF328684	AF328684 Homo sapi
C 35	182.4	5.9	55577	2	AC107915	AC107915 Homo sapi
C 36	182.4	5.9	117296	9	AC092865	AC092865 Homo sapi
C 37	179.4	5.8	582	6	AX456965	AX456965 Sequence
C 38	177.8	5.7	753	6	AX456967	AX456967 Sequence
C 39	177.6	5.7	110000	2	AC092450_2	Continuation (3 of
C 40	169.6	5.4	1013	6	AX456966	AX456966 Sequence
C 41	167.6	5.4	176924	2	AC111632	AC111632 Rattus no
C 42	166.2	5.3	125818	2	AC098495	AC098495 Rattus no
C 43	154.8	5.0	558	6	AX456968	AX456968 Sequence
C 44	154.8	5.0	1192	10	BC034893	BC034893 Mus muscu
C 45	148	4.8	55577	2	AC107915	AC107915 Homo sapi

ALIGNMENTS

RESULT 1  
AC092746/c  
LOCUS AC092746 109320 bp DNA linear PRI 12-JUN-2002  
DEFINITION Homo sapiens 12 BAC RP11-561P12 (Roswell Park Cancer Institute  
Human BAC Library) complete sequence.  
ACCESSION AC092746  
VERSION AC092746  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 109320)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,

Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, M., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinich, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseqed, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mathewney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R., Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (25-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (07-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 12, 2002 this sequence version replaced gi:21217374.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches

of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	Source
	Location/Qualifiers
	1..109320
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-561P12"
misc_feature	1..2004
	/note="overlaps bases 69129..71132 of clone AC092184"
	/function="clone overlap"
repeat_region	787..1014
	/rpt_family="MIR"
repeat_region	complement(1235..1452)
	/rpt_family="MIR"
repeat_region	1686..1789
	/rpt_family="L2"
repeat_region	1927..2209
	/rpt_family="AluSq"
repeat_region	2210..2238
	/rpt_family="(TAAAA)n"
repeat_region	2430..2741
	/rpt_family="AluJo"
repeat_region	2938..3085
	/rpt_family="L2"
repeat_region	complement(3086..3380)
	/rpt_family="AluSg"
repeat_region	3381..3665
	/rpt_family="L2"
repeat_region	3698..3940
	/rpt_family="AluJb"
repeat_region	4033..4297
	/rpt_family="AluSx"
repeat_region	4366..4397
	/rpt_family="(TTTA)n"
repeat_region	complement(4398..4689)
	/rpt_family="AluJb"
repeat_region	complement(4706..4993)
	/rpt_family="AluSx"
repeat_region	5253..5592
	/rpt_family="AluSx"
repeat_region	complement(5611..5872)
	/rpt_family="L2"
repeat_region	6027..6315
	/rpt_family="MLTIF"
repeat_region	complement(6371..6498)
	/rpt_family="MIR"
repeat_region	6703..7006
	/rpt_family="AluSx"
repeat_region	complement(7277..7397)
	/rpt_family="FLAM_C"







QY	634	GT	CAGATTTTGGCACCTAGGTGAGGCCAATCATCTCTGCAGAGCAATGTCCTCAATATGC	693
Db	626	GT	CAGGTTCTGGCACCCCATGAACCAATCTTCCAGAGAGCGGTGTGTTCAATAGTT	685
QY	694	TT	CTGAAACCTACAGATGGGGCTGGGAATGATTTATCTGTGAACCTAGAAGAAATTC	753
Db	686	T	ACTGAAATCTTGAATTTGGGCTTGAATGATGTTTCTGTGATAGTAAACACAATTC	745
QY	754	AT	ATGTGAGATGAATAAGATTTTACCTATGAGT	785
Db	746	AT	ATGTGAAATGAAGAAGATTTACCTATGAGT	777
RESULT 3				
AF240357				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
CDS				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches				
QY	40	GC	AGTTTGTCCCTGAGCTCTAGCTTCTTTAAATGAAGCTGAGTCTCTGGGCAACATCTTT	99
Db	39	GG	AGTTGATTTCTGAACCTCTGCCTCTTTTGACAGAGCCAGGTCCTCGATGCTGATTTT	97
QY	100	AG	GAGGAGAGGTACAAAAGTTCTTGGACCTTCTTCAACACAGGAGGCTGCATAATGATG	159
Db	98	--	GGACACAGATGCAAGAAACCCCT-GACCTTCTGAACATA--CACTCAACAATGGTG	151
QY	160	CA	AGACGACCAACCTCAAGATACAGAGAAAGAGCTGGTTGTCCTCGAGACTCTGGTCT	219
Db	152	C	AGGAAGACATATCCCA-----AGGAGAGGAGTCTCTGGACCCCTCAGACTCTGGTCA	205
QY	220	GT	GCGTGGGATTTCCATTCCTCAGTGTCTTGCCTTTCATTTGAGCTGTGTAGTAAC	279
Db	206	G	CTGCTGTGATTTCCATTTGAGTACCTTGTTCATTTGCGAGCTGTGTGGTGACT	265
QY	280	T	ACCATTTTACATATGTTGAAACCTGGCAAAAGGCTGTCTGAACCTACACTCATATCAT	339
Db	266	T	ACCAATTTTATTTATGGACACGCCAGTAGAAGACTATATGAACCTTCACACATACCA	325
QY	340	AG	TCTCACCTGCTTTCAGTGAAGGGACAAAGTGGCAG-----CCTGGGATGTTGGCCA	393
Db	326	AG	TCTCACCTGCTTTCAGTGAAGGGACTATGGTGCAGAAAAAATGCTGGGATGCTGCCA	385
QY	394	G	CTTCTTGGGAAGTCATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAGAGTTGG	453
Db	386	A	ATCACTGGAAGTCATTTTGGCTCCAGCTGCTACCTCATTTCTACCAAGGAGAACTTCT	445
QY	454	T	CTAAGAGTGAACAGAACTGTGTCAGATGGGAGCAGACATTTGGTTGTTTCAACACAG	513
Db	446	A	GCACAGTGAACAGAACTGTGTCAGATGGGCTCATCTGCTGGTGGTCAATACTAGAA	505
QY	514	GC	AGACAGCAATTTTCATTTGTCAGCAGCTGAATGAGTCATTTCTTATTTTCTGGGCTT	573
Db	506	G	CGGAGCAGAAATTTTCATCACCCAGCAGCTGAATGAGTCACCTTCTTCTGGGCTT	565
QY	574	T	CAGACCCACAAGTAAATAATTTGGCAATGGATTTGATTAAGACACACCTTTATGAA	633
Db	566	T	CGGATCCACAAGGTAAATGGCAATGGCAATGGATCGATGATCTCTCTTTCAGTCAAA	625
QY	634	G	TACATTTTGGCAGCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGCTTCAATAGTC	693
Db	626	G	TACAGTTCTGGCACCCCATGAACCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTT	685
QY	694	T	TCTGGAACCTACAGATGGGCTGGAAATGATGTTATCTGTGAAACTAGAAGGAATTC	753
Db	686	T	ACTGGAATCTTCGAAATGGGCTGGAAATCATGTTTCTGTGATGATTAACACAATTC	745
QY	754	AT	ATGTGAGATGAATAAGATTTTACCTATGAGT	785
Db	746	AT	ATCTGAAATGAAGAAGATTTTACCTATGAGT	777
RESULT 4				
BC023008				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REMARK				
COMMENT				

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 53 Row: P Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9910157.

FEATURES

source  
1..1254  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N-3"  
/clone="MGC:35901 IMAGE:5376247"  
/tissue\_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."  
/clone\_lib="NCI-CGAP\_Mam2"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
146..775  
/codon\_start=1  
/product="C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10"  
/protein\_id="AAH23008.1"  
/db\_xref="GI:18606438"  
/db\_xref="LocusID:56620"  
/translation="MVQEROSGKGVGCVTLRLWSAAVISMILLSTCFIASCVVVYQFI MDPSRLYELHTVSHSLTFCSEGTVMYSEKMGCCPNHKSFGSSCYLLSTKFNWST SEQNCVOMGHLVINTAEQONFTIOLNLSLYFLGLSDPQNGKQWIDDTFFSQN VRWFHEPNLPERCQSVIWNPSKMGVDFDCDSKHNSICEMKKIYL"

CDS

362 a 271 c 276 g 345 t  
BASE COUNT  
ORIGIN  
Query Match 12.2%; Score 379.2; DB 10; Length 1254;  
Best Local Similarity 73.8%; Pred. No. 1.2e-53;  
Matches 555; Conservative 0; Mismatches 178; Indels 19; Gaps 5;

QY 40 GCAGTTTGTCCCTGAGCTCTAGCTTCTTTAAATGAAGCTGAGTCTCTGGGCAACATCTTT 99  
DB 39 GGAAGTTGATCTGAACCTGGCTCTTTGACAGAAGCCAGTCCCTGAGTCGTATTTT- 97  
QY 100 AGGAGAGAGGTACAAAAGTTTCCGTGGACCTTCTCAACACAGAGGAGCCTGCATGATG 159  
DB 98 --GGAGACAGATGCAAGAAACCCCT-GACCTTCTCAACATA---CACCTCAACAATGGTG 151  
QY 160 CAAGAGCAGCAACCTCAAGTACAGAGAAAGAGGCTGGTTGTCCTCGAGACTCTGGTCT 219  
DB 152 CAGGAAGACAAATCCCA-----AGGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCA 205  
QY 220 GTGGCTGGGATTTCCATGCGACTCCTCAGTCTGCTTCATTGTGAGCTGTGTAGTAAC 279  
DB 206 GCTGCTGTGATTTCCATGTTACTCTGTAGTACCTGTTCATTGGCAGCTGTGTGACT 265  
QY 280 TACCATTTFACATATGGTGAACACTGGCAAAAGGCTGTCTGAACACACTACACTATATCA 339  
DB 266 TACCAATTTATTATGGACAGCCAGTGAAGACTATATGAACCTTCACACATACCATTC 325  
QY 340 AGTCTCACTGCTTCAGTGAAGGCAAGGTGCCAG-----CCTGGGATGTGGCCA 393  
DB 326 AGTCTCACTGCTTCAGTGAAGGAGTATGGTGTGAGAAAATGTGGGGAIGCTGCCCA 385  
QY 394 GCTTCTTGGAGTCAATTTGGTTCCAGTTGCTACTTTCATTTCAGTGAAGAAAGTTTGG 453  
DB 386 AATCACTGGAAGTCAATTTGGTCCAGCTGTACCTCAITTTCTACCAAGGAGAACTTCGG 445  
QY 454 TCTAAGATGAGCAGAACTGTGTGAGATGGAGACATATTTGGTTGTGTTTCAACACAGAA 513  
DB 446 AGCACAGGTGAGCAGAACTGTGTTCCAGATGGGGCTCATCTGGTGTGATCAATACTGAA 505  
QY 514 CGAGAGCAGAAATTTCAATCTCCAGCAGCTGAATGATCATTTCTTATTTTCTGGGGCTT 573  
DB 506 GCGGAGCAGAAATTTCAATCCAGCAGCTGAATGATGATCCTTTCTTACTTCTGGGCTT 565

QY 574 TCAGACCCACAGGTAATTAATTAATTTGGCAATGGATTGATAAGACACACCTTATGAGAAAAAT 633  
DB 566 TCGGATCCACAGGTAATGCAAAATGCAATGGATGGATGATGATCTCTTTCAGTCAAAAT 625  
QY 634 GTCAGATTTTGGCACCTAGGTAGGCCCCAATCATCTGCAGAGCAATGTCTTCAATATGTC 693  
DB 626 GTCAGTTCTTGGCAGCCCCCATGAACCAATCTTCCAGAAGAGGGTGTGTTTCAATATGTT 685  
QY 694 TTCTGAAACCTACAGGATGGGGCTGGAATGATGTTTATCTGTGAAACTAGAAGGAATCA 753  
DB 686 TACTGGAATCTTCGAATGGGCTGGAATGATGTTTCTGTGATAGTAACACAAATCA 745  
QY 754 ATATGTGAGATGAATAAGATTTACCTATGAGT 785  
DB 746 ATATGTGAAATGAAGAAGATTTACCTATGAGT 777  
RESULT 5  
AX014883 LOCUS 3708 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 81 from Patent WO9953040.  
ACCESSION AX014883  
VERSION AX014883.1 GI:10041150  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Human  
REFERENCE 1 (bases 1 to 3708)  
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.  
TITLE Human nucleic acid sequences from ovarian tumour tissue  
JOURNAL Patent: WO 9953040-A 81 21-OCT-1999;  
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)  
FEATURES  
Location/Qualifiers  
source 1..3708  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1036 a 818 c 671 g 1183 t  
ORIGIN  
Query Match 11.7%; Score 364.6; DB 6; Length 3708;  
Best Local Similarity 88.3%; Pred. No. 2.7e-51;  
Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;

```

|||||
Db 388 AAGTCTGAGTTCAAAAGTCATTTTTCCTCCACAAAGTTTCAACACTTAGCTAGAA 447
|||||
Qy 2906 CTTTTCAGTGTAACTTGCCT-AAAAGTTAAGACATATCTCGAGATCATATAAGTCAC 2964
|||||
Db 448 CTTTTCAGTGTAACTTGCCTAAAAGTTAAGACATATCTCGATATCATACACAGTCAC 507
|||||
Qy 2965 ATGATTTCTGATCTATCTCTGTGTTAATAACAAGATTTC 3007
|||||
Db 508 ATGATTTCTGATCTATCTGTGTTAATAATAAAGTCITTA 550
|||||

RESULT 6
AP004074
LOCUS Homo sapiens genomic DNA, chromosome 8q23, clone:KB1266F9, complete
DEFINITION
ACCESSION AP004074
VERSION AP004074.1 GI:15216353
KEYWORDS
SOURCE Homo sapiens cell_line:FLEB 14-14 DNA, clone_lib:Keio BAC library
clone:KB1266F9.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Shimizu,N. and Asakawa,S.
AUTHORS
TITLE Homo sapiens DNA chromosome 8 SEQUENCE
JOURNAL Published Only in Database (2001)
REFERENCE
2 (bases 1 to 156441)
AUTHORS Shimizu,N. and Asakawa,S.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2001) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
160-8582, Japan (E-mail:nshimizudb.med.keio.ac.jp,
Tel:81-3-3351-2370, Fax:81-3-3351-2370)
FEATURES
Source
1. 156441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1266F9"
/cell_line="FLEB 14-14"
/clone_lib="Keio BAC library"
560..581
/revidence=not_experimental
repeat_region /rpt_family="AT_rich"
2623..3033
/revidence=not_experimental
repeat_region /rpt_family="MSTB"
3626..3688
/revidence=not_experimental
repeat_region /rpt_family="(A)n"
5506..5537
/revidence=not_experimental
repeat_region /rpt_family="CT-rich"
7579..7632
/revidence=not_experimental
repeat_region /rpt_family="AT_rich"
complement(7956..8247)
/revidence=not_experimental
repeat_region /rpt_family="AluSx"
8697..8731
/revidence=not_experimental
repeat_region /rpt_family="(CAAA)n"
9452..9560
/revidence=not_experimental
repeat_region /rpt_family="L2"
10358..10385
/revidence=not_experimental
repeat_region /rpt_family="AT_rich"
10421..10441
/revidence=not_experimental
/revidence=not_experimental
/rpt_family="AT_rich"
complement(10443..10592)
/revidence=not_experimental
/rpt_family="FLAM_C"
11163..11185
/revidence=not_experimental
/rpt_family="(A)n"
11849..11876
/revidence=not_experimental
/rpt_family="AT_rich"
12699..12861
/revidence=not_experimental
/rpt_family="FAM"
13403..13423
/revidence=not_experimental
/rpt_family="AT_rich"
13609..13719
/revidence=not_experimental
/rpt_family="MIR"
13731..13758
/revidence=not_experimental
/rpt_family="AT_rich"
13766..14028
/revidence=not_experimental
/rpt_family="L1MC5"
14034..14260
/revidence=not_experimental
/rpt_family="Trigger1"
complement(14263..14407)
/revidence=not_experimental
/rpt_family="FLAM_A"
14409..14877
/revidence=not_experimental
/rpt_family="Trigger1"
14872..15595
/revidence=not_experimental
/rpt_family="Trigger1"
15637..15922
/revidence=not_experimental
/rpt_family="L1MC5"
17431..17469
/revidence=not_experimental
/rpt_family="AT_rich"
19504..19526
/revidence=not_experimental
/rpt_family="AT_rich"
complement(20688..21125)
/revidence=not_experimental
/rpt_family="L1MC1"
complement(22612..22919)
/revidence=not_experimental
/rpt_family="L1MD1"
complement(22922..23099)
/revidence=not_experimental
/rpt_family="AluJo"
complement(23100..23442)
/revidence=not_experimental
/rpt_family="L1MD1"
23442..24177
/revidence=not_experimental
/rpt_family="L1MC/D"
24178..24474
/revidence=not_experimental
/rpt_family="AluY"
24475..24685
/revidence=not_experimental
/rpt_family="L1MC/D"
25120..25255
/revidence=not_experimental
/rpt_family="FLAM_C"
complement(26608..26781)
/revidence=not_experimental

```

```

/rpt_family="MIR"
complement(26999..27171)
/evidence-not_experimental
/rpt_family="MER2"
27399..27410
/evidence-not_experimental
/rpt_family="AT_rich"
27411..27435
/evidence-not_experimental
/rpt_family="(CA)n"
29520..29564
/evidence-not_experimental
/rpt_family="G-rich"
30299..32109
/evidence-not_experimental
/rpt_family="LIMA4"
33154..33345
/evidence-not_experimental
/rpt_family="MIR"
33496..33574
/evidence-not_experimental
/rpt_family="CT-rich"
34340..34598
/evidence-not_experimental
/rpt_family="AluSc"
34711..34754
/evidence-not_experimental
/rpt_family="AT_rich"
35069..35113
/evidence-not_experimental
/rpt_family="GA-rich"
35622..35733
/evidence-not_experimental
/rpt_family="FLAM_A"
35943..35972
/evidence-not_experimental
/rpt_family="AT_rich"
complement(36516..36600)
/evidence-not_experimental
/rpt_family="MER5A"
36507..36687
/evidence-not_experimental
/rpt_family="T-rich"
36731..36961
/evidence-not_experimental
/rpt_family="MIR"
complement(36969..37228)
/evidence-not_experimental
/rpt_family="AluJb"
37373..37775
/evidence-not_experimental
/rpt_family="MUT1J"
38709..38978
/evidence-not_experimental
/rpt_family="LIMC5"
39053..39073
/evidence-not_experimental
/rpt_family="AT_rich"
39582..39719
/evidence-not_experimental
/rpt_family="L2"
complement(39720..40014)
/evidence-not_experimental
/rpt_family="LIMB5"
40015..40420
/evidence-not_experimental
/rpt_family="MSTP"
complement(40421..40462)
11.5%; Score 357.2; DB 9; Length 156441;
Best Local Similarity 88.3%; Pred. No. 2.5e-50;
Matches 454; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

QY 2509 TACAGCTTATACACAACTTTTATTAGAAAAGTTATATACACAGCATCAACTATTTT 2568
Db 798 TACAGCTTATACACAACTTTTATTAGAAAAGTTATATACATACATAGCATCAACTATTTT 857
QY 2569 CAAG-----AACCCCAATAGCAACAAACAAACCCAGACTAACAAAATGTGTACAAGA 2618
Db 858 CAAGAACAATATTAAACCCGATAGCAACAAACAAACCCAGACTAACAAAATGTGTACAAGA 917
QY 2619 AACTAATGACCTTTCTTAAATCAACATTCATTTATCTACATGCTCTTTTACAACAGG 2678
Db 918 AACTAATGACCTTTCTTAAATCAACATTCATTTATCTACATGCTCTTTTACAACAGG 977
QY 2679 GAAACCTCCATGGTTTACAGGCATGTCATATGAAAATAAAGCTGCAATAGC-TTTTAT 2737
Db 978 GAAACCTCCATGGTTTACAGGCATGTCATATGAAAATAAAGCTGCAATAGC-TTTTAT 1037
QY 2738 ACAATTATCGCTCTCAAGAAAATGAATCATTAAGCAGTAATTAGGAGTTTCAAAATTTA 2797
Db 1038 ACAATTATCGCTCTCAAGAAAATGAATCATTAAGCAGTAATTAGGAGTTTCAAAATTTA 1097
QY 2798 AACATTTCACGTAATTTTAAATATTGCTTTCATTAATTTTAAATTTTAAAGTCTGAG 2857
Db 1098 AACATTTCACGTAATTTTAAATATTGCTTTCATTAATTTTAAAGTCTGAG 1145
QY 2858 TTTCAAAAGTGA--TTTTCCTCCCAAGAGTGCCCAACTTAAGCTAGAGCTTTCAAGTG 2914
Db 1146 TTTCAAAAGTGATTTTTTTTCCCAAAAGTTTCAACACTTAAGCTAGAGCTTTCAAGTG 1205
QY 2915 TTAACCTTGGCCCT-AAAGTTAAGACATATCTGAGAATCATTAAGTACATGATTTCT 2973
Db 1206 TTAACCTTGGCCCTAAAGTTAAGACATATCTGAGAATCATTAAGTACATGATTTCT 1265
QY 2974 GATGCTATGCTCTGTTTAAATCAACAAAGATTTC 3007
Db 1266 GATGCTATGCTCTGTTTAAATCAACAAAGATTTC 1299

RESULT 7
AK097915/c 2209 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ40596 fis, clone THYM02010831, highly similar
DEFINITION to Mouse NCBP-29 mRNA for PW29.
ACCESSION AK097915
VERSION AK097915.1 GI:21757818
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens shymus cDNA to mRNA, clone_lib:THYM02
clone:THYM02010831.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,K., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 2209)
REFERENCE Isogai,T. and Yamamoto,J.
AUTHORS Direct Submission
TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and

```



```
Qy 2916 TAACTTTGCCCT-AAAGTTAAGACATATTCTGAGAATCAATAATAGTCACATGATTCTG 2974
|||||
Db 3239 TAACTTTGCCCTAAAGTTAAGACATATTCTGATGATCAATACACAGTCACATGATTCTG 3180
|||||
Qy 2975 ATGCTATCTGCTGTTAATAACAAAGATTTC 3007
|||||
Db 3179 ATGCTATCTGCTGTTAATAAAGTCTTTA 3147

RESULT 9
AC016729/c
LOCUS
DEFINITION AC016729 193100 bp DNA linear HTG 07-JUL-2000
Homo sapiens chromosome 8 clone RP11-367C15, WORKING DRAFT
SEQUENCE, 9 unordered pieces.
ACCESSION AC016729
VERSION AC016729.6 GI:8954332
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Jul 7, 2000 this sequence version replaced gi:7231020.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0367C15
----- Summary Statistics -----
Sequencing vector: M13; 78%
Sequencing vector: plasmid; 24%
Chemistry: Dye-primer ET; 76% of reads
Chemistry: Dye-terminator Big Dye; 24% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 186568 bases at least Q40
Consensus quality: 189080 bases at least Q30
Consensus quality: 190451 bases at least Q20
Insert size: 199000; agarose-fp
Insert size: 192300; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.86 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1597: contig of 1597 bp in length
* 1598 1697: gap of unknown length
* 1698 8306: contig of 6609 bp in length
* 8307 8406: gap of unknown length
* 8407 15335: contig of 6929 bp in length
* 15336 15435: gap of unknown length
* 15436 22120: contig of 6685 bp in length
* 22121 22220: gap of unknown length
* 22221 30889: contig of 8669 bp in length
* 30890 30989: gap of unknown length
* 30990 39394: contig of 8405 bp in length
* 39395 39494: gap of unknown length
```

```
* 39495 56109: contig of 16615 bp in length
* 56110 56209: gap of unknown length
* 56210 99245: contig of 43036 bp in length
* 99246 99345: gap of unknown length
* 99346 193100: contig of 93755 bp in length.
FEATURES
source
Location/Qualifiers
1..193100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-367C15"
1..1597
/misc_feature /note="assembly_name:Contig3"
1698..8306
/misc_feature /note="assembly_name:Contig4"
8407..15335
/misc_feature /note="assembly_name:Contig5"
15436..22120
/misc_feature /note="assembly_name:Contig6"
clone_end:T7
vector_side:right"
2221..30889
/misc_feature /note="assembly_name:Contig7"
30990..39394
/misc_feature /note="assembly_name:Contig8"
39495..56109
/misc_feature /note="assembly_name:Contig9"
56210..99245
/misc_feature /note="assembly_name:Contig10"
99346..193100
/misc_feature /note="assembly_name:Contig11"
clone_end:SP6
vector_side:left"
BASE COUNT 57926 a 37964 c 37237 g 59165 t 808 others
ORIGIN
Query Match 11.1%; Score 346.6; DB 2; Length 193100;
Best Local Similarity 87.1%; Pred. No. 1.4e-48;
Matches 447; Conservative 0; Mismatches 39; Indels 27; Gaps 5;
Qy 2509 TACAGCTTATACACAACTTTTATTAGAAAAGTTATACATAACACAGCATCAACTATTTT 2568
|||||
Db 22971 TACAGCTTATACACAACTTTTATTAGAAAAGTTATACATAACATAGCATCAACTATTTT 22912
|||||
Qy 2569 CAAG-----AACCCCAATAGCAACAAACCAACAGACTACAAATGTGTAAACAAG 2618
|||||
Db 22911 CAAGACAATATTAAACCCGATAGCAACAAACCAACAGACTACAAATGTGTAAACAAG 22852
|||||
Qy 2619 AACTAATGACCTTTCTAAATCAACATTCATTTATCTACAATGTCTATTACAAAACAGG 2678
|||||
Db 22851 AACTAATGACCTTTCTAAATCAACATTCATTTATCTACAATGTCTTTTACAAAACGGG 22792
|||||
Qy 2679 GAAACCTCCATGGTTTACAGGCATGTCATATTGAAAATAAGCTGCAATAGC-TTTTAT 2737
|||||
Db 22791 GAAACCTCCATGGTTTACAGGCATGTCATATTGAAATGTAAGCTGCAATAGCAATTTAT 22732
|||||
Qy 2738 ACAATTATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAAGAGTTCACAAATTTA 2797
|||||
Db 22731 ACAATTACCACCTCTCAAGAAAATGAATCATTAAGACAGTAATTAAGAGTTCACAAATTTA 22672
|||||
Qy 2798 AAACATTTCAAGTAATTTTAAATTTATTTGCTCTCAATAATTTTAAATTTATTTGAAGTCTGAG 2857
|||||
Db 22671 AAACATTTCAAGTAATTTTAAATTTATTTGCTCTCAATAATTTTAAATTTATTTGAGTCTGAG 22624
|||||
Qy 2858 TTTCAAAAGTGA---TTTTTCCCAACAAAGGTGCCAACACTTAAGCTAGAGCTTTTCAGTG 2914
|||||
Db 22623 TTTCAAAAGTGAATTTTTCCTCCCAACAAAGGTGCCAACACTTAAGCTAGAGCTTTTCAGTG 22564
|||||
Qy 2915 TTAACCTTGGCCCTAAAAGTAAAGACATATTTCTGAGATCATATAGTACATGATTTCTG 2974
|||||
Db 22563 TTAACCTTGGCCCTAAAAGTAAAGACATATTTCTGAGATCATATAGTACATGATTTCTG 22505
|||||
Qy 2975 ATGCTATCTGCTGTTAATAAACAAGATTTC 3007
```



||||| ATGCTATCTGCTGCTGCTTAATAAAGTCTTTA 22472

# RESULT 10

AC087350/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC087350 193364 bp DNA linear PRI 21-AUG-2002  
Homo sapiens chromosome 8, clone RP11-367C15, complete sequence.

AC087350

AC087350.3

GI:22267854

HTG

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 193364)

Birren, B., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 8, clone RP11-367C15

Unpublished

2 (bases 1 to 193364)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A.,

Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,

Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,

Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Gallagan, J.,

Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,

Lecoczky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P.,

Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,

McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,

Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,

Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,

Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,

Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 193364)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,

Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,

Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Gallagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,

Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 193364)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,

Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,

Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Gallagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 16, 2002 this sequence version replaced gi:16905274.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIDR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11698

Center clone name: 367\_C\_15

-----

FEATURES

source

1. 193364

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-367C15"

/clone\_lib="RP11-11 Human Male BAC"

290. 348

/rpt\_family="MIR"

349. 651

/rpt\_family="AluSq"

652. 730

/rpt\_family="MIR"

1721. 2290

/rpt\_family="L1MB5"

3577. 3711

/rpt\_family="MSTD"

3728. 3786

/rpt\_family="FRAM"

3812. 4016

/rpt\_family="MSTD"

complement(4135. 4438)

/rpt\_family="AluY"

4531. 4597

/rpt\_family="(TTCC)n"

5406. 6437

/rpt\_family="L1MEC"

6438. 6744

/rpt\_family="AluY"

6745. 7328

/rpt\_family="L1MEC"

8016. 8321

/rpt\_family="AluSx"

complement(8338. 8431)

/rpt\_family="L3"

complement(9625. 9835)

/rpt\_family="MLT1C"

complement(9836. 10138)

/rpt\_family="AluSx"

complement(10139. 10431)

/rpt\_family="MLT1C"

10441. 11704

/rpt\_family="L2"

12286. 12710

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

```

/rpt_family="LIPAL6"
12715..12735
/rpt_family="TAAA)n"
12806..12829
/rpt_family="HAL1"
12830..13136
/rpt_family="AluSx"
13137..13265
/rpt_family="HAL1"
13791..13843
/rpt_family="L2"
complement(14375..14530)
/rpt_family="MIR3"
15285..16638
/rpt_family="PTR5"
complement(16805..17197)
/rpt_family="MLTIL"
complement(17353..17462)
/rpt_family="MIR"
complement(17890..18020)
/rpt_family="MIR3"
18233..18405
/rpt_family="MIR"
complement(18743..18834)
/rpt_family="MIR"
18857..18915
/rpt_family="MIR"
complement(21125..21435)
/rpt_family="AluSx"
21499..21528
/rpt_family="AT_rich"
complement(22565..22867)
/rpt_family="AluY"
23040..23060
/rpt_family="AT_rich"
24278..24530
/rpt_family="L2"
24532..24562
/rpt_family="(TTAA)n"
complement(24678..24746)
/rpt_family="Charlie4"
complement(25284..25429)
/rpt_family="MIR"
25489..25526
/rpt_family="5S"
complement(25586..25737)
/rpt_family="MIR3"
complement(25941..26235)
/rpt_family="AluSq"
27079..27100
/rpt_family="(CCG)n"
28055..28463
/rpt_family="MSTA"
28644..28671
/rpt_family="(TTCA)n"
complement(28724..28949)
/rpt_family="MIR"
complement(29453..29575)
/rpt_family="L2"
complement(29580..29674)

Query Match 11.1% Score 346.6; DB 9; Length 193364;
Best Local Similarity 87.1%; Pred. No. 1.4e-48;
Matches 447; Conservative 0; Mismatches 39; Indels 27; Gaps 5;

QY 2509 TACAGCTTATACACAACTTTATTAGAAAGTTATACATACACAGCAGTCAACTATTTT 2568
|||||
Db 119652 TACAGCTTATACACAACTTTATTAGAAAGTTATACATACACAGCAGTCAACTATTTT 119593
|||||

QY 2569 CAAG-----AACCCCAATAGCAACAAACAAACAGACTAACAAAAATGGTAAACAAGA 2618
|||||
Db 119592 CAAGCAATATTAAACCCGATAGCAACAAACAAACAGACTAACAAAAATGGTAAACAAGA 119533
|||||

```

QY 2619 AACTAATGACCTTTCTTAAATCAAACTCAATTTATCTACATGCTATTTTACAAACAGG 2678  
 |||||||  
 Db 119532 AACTAATGACCTTTCTTAAATCAAACTCAATTTATCTACATGCTATTTTACAAACAGG 119473  
 |||||||  
 QY 2679 GAAAACTCCATGTTTACAGGATGTCATATGAAAATAAAGCTGCAATAGC-TTTTAT 2737  
 |||||||  
 Db 119472 GAAAACTCCATGTTTACAGGATGTCATATGAAAATAAAGCTGCAATAGC-TTTTAT 119413  
 |||||||  
 QY 2738 ACAATTATCGTCTCAAGAAATGAATCAATTAAGACATTAATAGGAGTTTCAAAATTA 2797  
 |||||||  
 Db 119412 ACAATTATCGTCTCAAGAAATGAATCAATTAAGACATTAATAGGAGTTTCAAAATTA 119353  
 |||||||  
 QY 2798 AACTATTCCAGTAAATTTAAATATTTGTTTCAATAATTTTAAATATTGAAGTCTGAG 2857  
 |||||||  
 Db 119352 AACTATTCCAGTAAATTTAAATATTTGTTTCAATAATTTTAAATATTGAAGTCTGAG 119305  
 |||||||  
 QY 2858 TTTCAAAAGTGA--TTTTCCTCCCAAGAGTGCCAACTTAAGCTAGAGCTTTCAGTG 2914  
 |||||||  
 Db 119304 TTTCAAAAGTGAATTTTTCCTCCCAAGAGTTTCAACACTTAAGCTAGAGCTTTCAGTG 119245  
 |||||||  
 QY 2915 TTAACCTTCCCTCAAAAGTTAAGACATATCTGAGAATCATATAGTCACATGATTTCTG 2974  
 |||||||  
 Db 119244 TTAACCTTCCCTCAAAAGTTA-AGACATTTCTGATAATCATACAGTCACATGATTTCTG 119186  
 |||||||  
 QY 2975 ATGCTATCTGCTCTGTTAATAACAAGATTTC 3007  
 |||||||  
 Db 119185 ATGCTATCTGCTCTGTTAATAACAAGTCTTTA 119153  
 |||||||

RESULT 11  
 HSU85A3  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone U85A3 on chromosome X contains CSF2  
 (cleavage stimulation factor, subunit 2, 64kD) pseudogene similar  
 to T-cell cyclophilin, pseudogene similar to S.pombe rad21, STSS  
 and GSSs, complete sequence.  
 278021  
 278021.1 GI:1483514  
 HTG: CSF2.  
 SOURCE  
 Homo sapiens  
 ORGANISM  
 Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 37652)  
 Odell,C.  
 AUTHORS  
 Direct Submission  
 TITLE  
 Submitted (31-JUL-1996) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 JOURNAL  
 requests: clonerequest@sanger.ac.uk  
 COMMENT  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated human repeat sequence elements (e.g. Alu). Where the  
 sequence is ambiguous, there is an annotation using the 'unsure'  
 feature key.  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome X, constructed by the Sanger Centre Chromosome X  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/ChrX  
 U85A3 is from the Lawrence Livermore National Laboratory  
 flow-sorted X chromosome cosmid library LLOXNC01.  
 FEATURES  
 Location/Qualifiers  
 1..37652  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /clone="LLOXNC01-85A3"

```

/clone_lib="LLOXNC01"
228..285
/notes="Alu repeat: matches 247..304 of consensus"
/evidence=not_experimental
gene
/complement(join(673..842,1488..1566,2619..2676))
/mRNA
/genes="CSTF2"
/complement(join(<673..842,1488..1566,2619..2676))
/notes="match: M85085"
/evidence=not_experimental
CDS
/complement(join(<673..842,1488..1566,2619..2676))
/genes="CSTF2"
/notes="match: protein P33240"
/codon_start=1
/evidence=not_experimental
/product="CSTF2 (cleavage stimulation factor, subunit 2, 64kb)"
/protein_id="CAB41262.1"
/db_xref="GI:4678893"
/translation="MAGLTVPQADVRLSRVFGNIPYEATBQLKIDIFSEVGPVVS
FRLYVDRETKPGYGCYQDETALSGMRNLNGREFSGRALRVDAASEKNKEELK
983..1281
/notes="AluX repeat: matches 1..300 of consensus"
/evidence=not_experimental
repeat_region
1324..1369
/notes="L1MB8 repeat: matches 6126..6173 of consensus"
/evidence=not_experimental
misc_feature
/complement(1919..2597)
/genes="CSTF2"
repeat_region
3581..3870
/notes="match: GSS B68800 clone 2023F10"
/evidence=not_experimental
repeat_region
4175..4354
/notes="MER5A repeat: matches 1..178 of consensus"
/evidence=not_experimental
repeat_region
4787..5081
/notes="MLT1E repeat: matches 264..568 of consensus"
/evidence=not_experimental
repeat_region
5208..5278
/notes="MLT2FA repeat: matches 118..198 of consensus"
/evidence=not_experimental
repeat_region
5391..5513
/notes="MIR repeat: matches 11..144 of consensus"
/evidence=not_experimental
repeat_region
7362..7921
/notes="L1MB3 repeat: matches 5614..6182 of consensus"
/evidence=not_experimental
repeat_region
7922..8420
/notes="MLT2CA repeat: matches 1..503 of consensus"
/evidence=not_experimental
repeat_region
8422..9916
/notes="L1MB3 repeat: matches 4109..5617 of consensus"
/evidence=not_experimental
repeat_region
9914..9994
/notes="L1MB4 repeat: matches 1853..1933 of consensus"
/evidence=not_experimental
repeat_region
10004..10872
/notes="L1PBA repeat: matches 1355..412 of consensus"
/evidence=not_experimental
repeat_region
10860..10971
/notes="L1PBA repeat: matches 279..160 of consensus"
/evidence=not_experimental
repeat_region
10975..11231
/notes="L1PBA repeat: matches 1058..3300 of consensus"
/evidence=not_experimental
repeat_region
13232..13536
/notes="AluY repeat: matches 1..304 of consensus"
/evidence=not_experimental
repeat_region
13537..14462
/notes="L1PBI repeat: matches 3299..4224 of consensus"
/evidence=not_experimental
14464..14947
/notes="L1R10A repeat: matches 1..548 of consensus"
/evidence=not_experimental
repeat_region
14973..16855
/notes="L1PBI repeat: matches 4234..6151 of consensus"
/evidence=not_experimental
repeat_region
16876..16998
/notes="L1M4C repeat: matches 1712..1835 of consensus"
/evidence=not_experimental
repeat_region
17090..17924
/notes="L1M4C repeat: matches 846..1603 of consensus"
/evidence=not_experimental
repeat_region
18176..18298
/notes="L1M3D repeat: matches 495..618 of consensus"
/evidence=not_experimental
repeat_region
18299..18460
/notes="L1M4 repeat: matches 79..236 of consensus"
/evidence=not_experimental
misc_feature
18501..18903
/notes="match: STS G23519"
/complement(19640..20032)
/misc_feature
/notes="match: STS G11279"
/complement(20033..21918)
gene
/genes="CU85A3.3"
CDS
/complement(20033..21918)
/genes="CU85A3.3"
/notes="match: D38551"
/pseudo
/codon_start=1
/evidence=not_experimental
/product="CU85A3.3 (similar to S.pombe rad21)"
20346..20787
/notes="match: STS G30625"
/misc_feature
20368..20616
/notes="match: STS G29205"
22030..22270
/notes="MER4C repeat: matches 143..447 of consensus"
/evidence=not_experimental
gene
22311..22799
/genes="CU85A3.2"
CDS
22311..22799
/genes="CU85A3.2"
/notes="match: P05092"
match: Y00052"
/pseudo
/codon_start=1
/evidence=not_experimental
/product="CU85A3.2 (similar to T-cell cyclophilin)"
23010..23320
/notes="AluSp repeat: matches 1..311 of consensus"
/evidence=not_experimental
repeat_region
23375..23691
/notes="MER4A repeat: matches 3..337 of consensus"
/evidence=not_experimental
repeat_region
23695..23867
/notes="MER4-internal repeat: matches 6450..6598 of consensus"
/evidence=not_experimental
repeat_region
23925..23982
/notes="29 copies 2 mer aa 76% conserved"
/evidence=not_experimental
repeat_region
24322..24501
/notes="HERV1L18 repeat: matches 5621..5801 of consensus"
/evidence=not_experimental
repeat_region
26573..26872
/notes="AluSq repeat: matches 1..301 of consensus"
/evidence=not_experimental
repeat_region
27271..27604
/notes="AluYa8 repeat: matches 1..307 of consensus"
/evidence=not_experimental
repeat_region
27645..27906
/notes="HUERS-P3 repeat: matches 4101..5118 of consensus"
/evidence=not_experimental

```



```

variation
6020. .6022
/notes="clone U85A3; GAC in this entry; substitution"
/replaces="ggc"
6435. .6584
/notes="LRR7 repeat: matches 448. .301 of consensus"
complement(6465. .11676)
/notes="remnants of endogenous retrovirus; partial homology
to viral RTVL-H, Phospholipase A2 homologue, GAG core shell
proteins, envelope (coat) proteins nucleic acid binding
proteins, POL polyprotein, reverse transcriptase,
ribonuclease H; matches many proteins, ESTs, cDNAs,
genomic DNAs and STSS"
7297. .7299
/notes="clone U85A3; AAA in this entry; substitution"
/replaces="ata"
8284. .8286
/notes="clone U85A3; GCA in this entry; substitution"
/replaces="gga"
8570. .8572
/notes="clone U85A3; GGA in this entry; substitution"
/replaces="gaa"
8622. .8624
/notes="clone U85A3; TAG in this entry; substitution"
/replaces="tgg"
9488. .9490
/notes="clone U85A3; AAT in this entry; substitution"
/replaces="agt"
11025. .11027
/notes="clone U85A3; TCT in this entry; substitution"
/replaces="ttt"
11321. .11620
/notes="LRR7 repeat: matches 448. .134 of consensus"
11572. .11681
/notes="LRR7 repeat: matches 131. .1 of consensus"
13573. .18746
/notes="remnants of endogenous retrovirus; partial homology
to viral envelope (coat) proteins, POL polyprotein,
reverse transcriptase, ribonuclease H; matches many
proteins, ESTs, cDNAs and genomic DNAs"
13669. .13671
/notes="clone U85A3; GCT in this entry; substitution"
/replaces="gtt"
13986. .13988
/notes="clone U85A3; TAT in this entry; substitution"
/replaces="tgt"
15110. .15112
/notes="clone U85A3; CCG in this entry; substitution"
/replaces="ctg"
15318. .15640
/notes="AluY repeat: matches 1. .299 of consensus"
16051. .16349
/notes="AluSg repeat: matches 2. .301 of consensus"
16632. .16634
/notes="clone U85A3; ATG in this entry; substitution"
/replaces="acg"
17593. .17595
/notes="clone U85A3; AAA in this entry; substitution"
/replaces="aga"
17984. .17986
/notes="clone U85A3; TTC in this entry; substitution"
/replaces="tgc"
18981. .18982
/notes="clone U85A3; CT in this entry; deletion"
/replaces="ctt"
19001. .19117
/notes="3 copies of 39 mer 82 % conserved"
19146. .19148
/notes="clone U85A3; ATC in this entry; substitution"
/replaces="agc"
19225. .19541
/notes="MER4A repeat: matches 3. .337 of consensus"
19439. .19441
/notes="clone U85A3; ATC in this entry; substitution"

```

```

repeat_region
variation
gene
mRNA

/replace="acc"
19800. .19902
/notes="AluSp repeat: matches 303. .1 of consensus"
19607. .19608
/notes="clone U85A3; AC in this entry; deletion"
/replace="atttac"
complement(19787. .20634)
/gene="dJ347M6.1"
complement(<19787. .>20619)
/gene="dJ347M6.1"
/product="Cyclophilin A (Cyclosporin A-binding protein,
pPIASE, peptidyl-prolyl cis-trans isomerase A, Spi8,
Rotamase) like protein"
/notes="match: cDNAs M25637 M19533 Y00052 X52803 X17105;
match: ESTs AA075585 AA299048 F04720 AA362886 AA488379
AA486598 AA311324 AA131423 T70141 AA307700 AA533008
AA513009 AA180415 AA314272 AA081118 AA300395 AA505262
D57979 AA353507 AA581820 AA126791 AA401598 AA335880
AA573841 AA171787 AA582439 AA299336 D55800 HA3341 F18530
D55283 AA315532 AA504064 AA182044 AA191632 AA302188
AA342661 AA191634 AA220938 AA126843 AA532728 AA488511
AA088390 AA307477 AA412697 AA576164 AA342786 AA147099
AA155677 AA552685 AA466693 R87887 W79891 AA101162 AA314411
AA321796 AA483724 AA373538 AA356609 AA147872 AA531206
AA070523 AA305622 AA307522 AA554115 AA346116 T69131
AA148353 AA522575 AA306872 AA308538 W87520 AA173745
AA398530 AA484804 W87521 AA157745 AA528078 AA385581
AA146938 AA313332 AA302560 AA328140 AA191136 AA306921
AA370383 AA181295 AA328545 AA321796 AA483724 AA373538
AA356609 AA147872"
/pseudo
/evidence=not_experimental
2009. .20011
/notes="clone U85A3; AGA in this entry; substitution"
/replace="aca"
complement(20116. .20634)
/gene="dJ347M6.1"

Query Match      10.2%; Score 317.6; DB 9; Length 56583;
Best Local Similarity 83.3%; Pred. No. 1.le-43;
Matches 418; Conservative 0; Mismatches 59; Indels 25; Gaps 4;

Qy 2504 TTTTTCAGCGTTATACACAACTTTTATTAGAAAGTTATACATAACACAGCATCAACT 2563
Db 24425 TTGTTTCAACTTATACACAACTTTTATTAGAGAGTTATACATAACATGATCAACT 24366

Qy 2564 ATTTTCAGAAC-----CCATAAGCAACAAACACAGACTAACAAATGTGTAAAC 2614
Db 24365 ATTTTCAAGAACATATTAAACCCCAATTAAGCAAAAAACACAGACTAACAAATGTGTAAAC 24306

Qy 2615 AAGAACTTAAGCCTTCTTAAATCAACATTCGAATTTATCTACAATGCTCTATTACAAA 2674
Db 24305 AAGAAACCAATGACCTTCTTAAATCAACATTCGAATTTATCTACAATGCTCTATTACAAA 24246

Qy 2675 CAGGAAAACCTCCATGGTTTACAGGCATGTCATATTGAAAAATAAAGCTGCAATAGCTTTT 2734
Db 24245 CAGGAAAACCTCCATGGTTTACAGGCACAGCATATTGAAATAAATCTGCAATAGCAATT 24186

Qy 2735 TATACAATTATCGCTCTCAAGAAATGAATCATTAAGACAGTAATAGGAGTTCACAAAT 2794
Db 24185 TTTTATGGTTAAACCTCTCAGAAACGGGAATCATTAACAAACAGTAAT---GAGTTTCACAAAT 24129

Qy 2795 TTAACAACTTTCACGTAATTTTAAATTTATTCCTCAATAATTTAAATTTTAAATTTTCAAGTCT 2854
Db 24128 TTAACAACTTTCACATAATTTTAAATTTATTCAGT-----ATACACTGAAGTCT 24081

Qy 2855 GAGTTTCAAAAGTGATTTTTTCCCAAAAGTGCCAACTTAAAGCTAGAGCTTTTCAGTG 2914
Db 24080 GTTTCAAAAGTGATTTTTTCCCAACGAAAGTGTCACACTTAAAGCTAGAGCTTTTCAGTG 24021

Qy 2915 TTAACCTTGCCCT-AAAAGTTAAGACATATTCTCAGATCATATAATAGTCACATGATTTCT 2973
Db 24020 TTAATTTTGCCCTAAAAAGTTAAGACATATTCTGATATAATAGTCACATGATTTCT 23961

```



Best Local Similarity 64.1%; Pred. No. 8.9e-38; Matches 488; Conservative 0; Mismatches 247; Indels 26; Gaps 3;									
Qy	59	TAGCTTCTTAAATGAAGCTGAGTCTCTGGCCACATCTTTAGGACAGAGGTACAAAAG	118						
Db	383	TGGCTTGTTGAAAGACCACACCCGAAAGTAACATCTTTGGAGAAAGGTATACAAG	442						
Qy	119	GTTCCTGGACC-----TTCTCAACACAGGGGCTGTGCATATGATGCAAGA	164						
Db	443	CTTCTGCACCCACCTGATAGAGGAAGTCCAAGGGTCTGGCACACACATGGTGCCTGA	502						
Qy	165	GCAGCAACCTCAAGTACAGAGAAAGAGGCTGGTTGTCCC---TGAGACTCTCGGTGTG	221						
Db	503	AGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGGTTCAGTGAAGGCTCTGGTCCAT	562						
Qy	222	GGCTGGGATTTCCATTTGCATCTCCTCAGCTGTCTGCTTCATTTGAGCTGTGTAGTAAC	281						
Db	563	GGCAGTGTATCCATCTTGTCTCCTCAGTGTCTGTTTCACTGTGAGTCTGTGTGGCTCA	622						
Qy	282	CCATTTTACATATGGTGAAGCTGGCAAAAGGCTGTCTGAAC	332						
Db	623	CAATTTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGACAGTATCAACAG	682						
Qy	333	TCATTCAAGTCTACCTGCTTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCC	392						
Db	683	TCATCCAAGCCTGACCTGCGTCTATGGAAGGAAGGACATAGAAGATTTGGAGCTGCTGCC	742						
Qy	393	AGCTTCTTGAAGTCATTTGGTTCACGTCTCTACTTTCATTTCCAGTGAAGAGAGGTTG	452						
Db	743	AACCCCTTGGACTTCATTTTCAGTCTAGTGTCTACTTTATTTTACTGGGATGCAATCTTG	802						
Qy	453	GTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTGTCTGTCAACACAG	512						
Db	803	GACTAAGAGTCAAAAGAACTGTCTGTGATGGGGCTGTACTGTGTGATCAACACCAG	862						
Qy	513	ACGACAGCAAAATTCATTTGTCAGCAGCTGAATGAGTCAATTTCTTTATTTCTGGGGCT	572						
Db	863	GGAAGAACAGGATTTTCATCTATCAGAACTCTGAAAGAAATTTCTTTATTTCTGGGGCT	922						
Qy	573	TTACAGCCCAACAGTAATAATATTTGGCAATGGATTGATAACACACCTTATCAGAAAA	632						
Db	923	GTGAGATCCAGGGGTGCGGCACATTTGGCAATGGTTGACCAGACCATACAAATGAAA	982						
Qy	633	TGTCAGATTTTGGACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTGCTTCAATAG	692						
Db	983	TGTCACATTTCTGGCACTCAGGTGAACCCAAATAACCTTGATGAGCGTTGTCCGATATA	1042						
Qy	693	CTTCTGAAACCTACAGGATGGGCTGGGAATGATGTTATCTGTGAACACTAGAAGGAAT	752						
Db	1043	TTTCCGTTCTTCAGAGAAATGGGGCTGGAATGACATTTCACTGTATGTACCTCAGAAG	1102						
Qy	753	AATATGTGAGATGAATAAGATTTACCTATGAGTAGAGCTT	793						
Db	1103	ANTTTCAGATGAAGAGATCTACATATAAATGAAATATT	1143						
RESULT 15 AF325459 LOCUS Homo sapiens dendritic lectin (CLECSF11) mRNA, complete cds, DEFINITION alternatively spliced. ACCESSION AF325459 VERSION AF325459.1 GI:17225336 KEYWORDS SOURCE Homo sapiens. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 851) Arce, I., Roda-Navarro, P., Montoya, M.C., Hernanz-Ruiz, E., AUTHORS Puig-Kroger, A. and Fernandez-Ruiz, E. TITLE Molecular and genomic characterization of human DLEC, a novel member of the C-type lectin receptor gene family preferentially									
expressed on monocyte-derived dendritic cells Eur. J. Immunol. 31 (9), 2733-2740 (2001) 21426806 MEDLINE PUBMED 11536172 REFERENCE 2 (bases 1 to 851) Arce, I., Hernanz, P. and Fernandez-Ruiz, E. AUTHORS Direct Submission TITLE Submitted (01-DEC-2000) Biologia Molecular, Hospital de la JOURNAL Princesa, Diego de Leon, 62, Madrid 28006, Spain FEATURES Location/Qualifiers source 1..851 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="12" /map="12p12.3-pl3.2" 1..851 /gene="CLECSF11" 27..668 /note="C-type lectin receptor; DLEC; alternatively spliced" /codon_start=1 /product="dendritic lectin" /protein_id="AAL37358.1" /db_xref="GI:17225337" /translation="MVPEEPQDREKGLWVFKVMSAVVSIILLVSCPTVSSVVP NPMYSKTVKRLSKLREYOYHPSLTVMEGKDIEDMCCPTPTWTSFSSCYFISTGMO SWTKSQKNCVMGADLVVINTREBODFLIQLKRNSSYFLGLSDPGRGRHWQVDDTP YNEVNTFWHSGEPNNLDERCAIINFRSEEWGNDIHCHVPKRSICKMKKIYI"									
BASE COUNT 242 a 169 c 192 g 248 t ORIGIN									
Query Match 9.0%; Score 280.6; DB 9; Length 851; Best Local Similarity 65.8%; Pred. No. 3.2e-37; Matches 444; Conservative 0; Mismatches 219; Indels 12; Gaps 2;									
Qy	131	TCCTCAACACAGGAGGCTGCATAATGATGCAAGACGACGCAACCTCAAAGTACAGAGAAA	190						
Db	4	TCCAAAGGGTGTGGCCACACACAATGGTGCCTGAAGAGAGCGCTCAAGACCAGAGAAAG	63						
Qy	191	GAGGCTGGTGTGCC---TGAGACTCTGGTCTGTGGCTGGGATTTCCATTTGCACCTCTCA	247						
Db	64	GACTCTGGTGGTTCAGGTTGAAGGCTGCTGCCATGGCAGTCGTATCCATCTTGTCTCTCA	123						
Qy	248	GTGCTTGTCTCATTTGAGCTGTGTAGTAACATTACCATTTTACATATGGTGAACCTGCA	307						
Db	124	GTGCTGTGTTTCTACTGTGAGTCTGTGGTGCCCTCACAAATTTTATGTATAGCAAAACTGT	183						
Qy	308	AAAGGCTGTCTCAACTA-----CACTCATATCATTTCAAGTCTCACTCTGTTCACTG	358						
Db	184	AGAGGCTGTCCAAGTTACGAGAGTATCAACAGTATCATCCAAGCCTGACCTGCGTCATGG	243						
Qy	359	AAGGACAAAGGTGCCAGCCTGGGATGTTGCCAGCTTCTTGGAGTCAATTTGTTGCCA	418						
Db	244	AAGGAAAGGACATAGAAGATTGGAGCTGCTGCCAACCCCTTGGACTTCATTTCACTAGTCA	303						
Qy	419	GTGCTACTTTCATTTCCACTGAAGAGAGGTTGGTCTTAAGAGTCAAGCAGAACTGTGTG	478						
Db	304	GTGCTACTTTTATTTCTACTGGGATGCAATCTTGGACTGAAGAGTCAAAAGAACTGTTCTG	363						
Qy	479	AGATGGGAGCACATTTGGTGTGTTTCAACACAGAGAGCAGAGAAATTTTCATTTGCCAGC	538						
Db	364	TGATGGGGCTGATCTGGTGGTATCAACACAGGAGAGAACAGGATTTTCATTCATTCAGA	423						
Qy	539	AGCTCAATGAGTCAATTTTCTTTCTGGGCTTTTCAACCCCAAGTCAATATAAT	598						
Db	424	ATCTCAAAAGAAATCTTCTTATTTTCTGGGCTGTCTCAGATCCAGGGGTGGCGACATT	483						
Qy	599	GGCAATGATGATGAAGACACCTTATGAGAAAATGTGAGATTTTGGCACCTAGGTGAGC	658						
Db	484	GGCAATGGGTTGACCAGACACCATACAAATGAAATGTCCACATTTCTGGCACTCAGGTGAC	543						
Qy	659	CCAATCATTTCTGCAGAGCAATGTGTCTTCAATAGTCTTCTGGAAACCTACAGATGGGCT	718						

```
Db 544 CCAATAACCTTGAAGCGTTGCGATAATAAATTCGGTTCTTTCAGAAGAATGGGCT 603
QY 719 GGAATGATCTTATCTGTGAACTAGAGGAATTCATATGTGAGATGAATAAGATTACC 778
Db 604 GGAATGACATTCAGTGTATGTACCTCAGAGTCAATTTGCAAGATGAAGAAGATCTACA 663
QY 779 TATGAGTAGAAGCTT 793
Db 664 TATAATGAAATATT 678
```

Search completed: February 17, 2003, 11:03:43  
Job time : 6410.63 secs





```

1  RESULT 2
2  US-08-772-440-20
3  ; Sequence 20, Application US/08772440
4  ; Patent No. 6046158
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Ariizumi, Kiyoshi
7  ; APPLICANT: Takashima, Akira
8  ; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
9  ; TITLE OF INVENTION: LECTINS, DEXTIN-1 AND DEXTIN-2; COMPOSITIONS AND USES
10 ; TITLE OF INVENTION: THEREOF
11 ; NUMBER OF SEQUENCES: 42
12 ; CORRESPONDENCE ADDRESS:
13 ; STREET: P.O. Box 4433
14 ; CITY: Houston
15 ; STATE: Texas
16 ; COUNTRY: USA
17 ; ZIP: 77210
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: Patentin Release #1.0, Version #1.30
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/08/772,440
25 ; FILING DATE: CONCURRENTLY HERewith
26 ; CLASSIFICATION: 435
27 ; ATTORNEY/AGENT INFORMATION:
28 ; NAME: Parker, David L.
29 ; REGISTRATION NUMBER: 32,165
30 ; REFERENCE/DOCKET NUMBER: UTXD:493
31 ; TELECOMMUNICATION INFORMATION:

```

```

RESULT 3
US-08-772-440-22
; Sequence 22, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```











```
; NAME/KEY: modified_base
; LOCATION: 3405..6871
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "S = C or G"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3457..9998
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "W = A or T"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3595..9999
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "Y = C or T"
US-08-772-440-33
```

```
Query Match 8.3%; Score 52; DB 3; Length 10409;
Best Local Similarity 80.3%; Pred. No. 2.1e-06;
Matches 61; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 294 GGTTCGTCTAGAGTGAGCAGAACTGTGTGAGATGGGAGCACATTTGGTTGTTCAA 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8777 GGCTTGGACACACGTGAGCAGAACTGTGTTCAGATGGGGGCTCATCTGGTGTGATCAA 8718

QY 354 CACAGAAGCAGACGAG 369
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8717 TACTGAAGCGGAGCAG 8702
```

```
RESULT 14
US-08-650-578-1
; Sequence 1, Application US/08650578
; Patent No. 5811284
; GENERAL INFORMATION:
; APPLICANT: Chang, Chiwen
; APPLICANT: Aramburu Beltran, Jose
; APPLICANT: Lopez-Botet, Miguel
; APPLICANT: Phillips Jr., Joseph H.
; APPLICANT: Lanier, Lewis L.
; TITLE OF INVENTION: Purified Mammalian NK Antigens and
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,578
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,339
; FILING DATE: 29-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 122..661
US-08-650-578-1

Query Match 7.0%; Score 44.2; DB 1; Length 871;
Best Local Similarity 53.1%; Pred. No. 0.00019;
Matches 94; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 201 CAGTGAAGGACAAAGTGCCAGCCTGGGATGTTGCCAGACTTCTTGGAAAGTCATTGG 260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 CATAGAACTCCAGAAAGACTCTGACTCTGTCTTGTGCAAGAAAAATGGGTGGGTACCG 327

QY 261 TTCCAGTTGCTACTTCTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGTGAGCAGAACTG 320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 GTGCAACTGTACTTCAATTTCCAGTGAACAGAAACTTGAACGAAGTCGGCATCTCTG 387

QY 321 TGTTCAGATGGGACACATTTGGTTGTTCACACAGAAGCAGACAGAGCAATTTCAAT 377
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 TGCTTCTCAGAAAAATCCAGCCTGCTTCAGCTTCAAAACACAGATGAATGGATTTTAT 444

RESULT 15
US-09-222-575-29/c
; Sequence 29, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, JIangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Human
US-09-222-575-29

Query Match 6.4%; Score 40.2; DB 4; Length 301;
Best Local Similarity 53.1%; Pred. No. 0.0022;
Matches 112; Conservative 0; Mismatches 93; Indels 6; Gaps 1;

QY 408 TTTTCGGGGCTTTCAGACCCCAAGTAATAATAATTGGCAATGGATTGATAAGACACC 467
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTTTATTTGGACTGTCAGACCAGGTTGTCGAGGGTCAATGGCAATGGGTGGACGCACACC 242

QY 468 TTATGACAAAAATGTCAGATTTTGGCACCTAGCTAGCCCAATCTCTGCA-----GA 521
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGGCCCAACAACATAGCTACCTGGA 182

QY 522 GCAATGTGCTTCAATAGTCTTCTGGAAACCTACAGGATGGGCTGGAATCATGTTATCTG 581
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GGACTGTGCCACCATGAGAGACTCTTCAACACCCCAAGGCAAAATTTGCAATCATGTAACCTG 122

QY 582 TGAAGTGAAGGAATTCATATATGTGAGATG 612
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 TTTCCCTCAATTATTTTCGGGATTTGTGAAATG 91
```

Search completed: February 17, 2003, 12:10:44  
Job time : 29.425 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:44:21 ; Search time 90.6728 Seconds  
(without alignments)  
15572.506 Million cell updates/sec

Title: US-09-766-511B-52

Perfect score: 627

Sequence: 1 atgatgcagagcagcaacc.....agatgaataagatttaacctta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	627	100.0	3114	22	Human TANGO 405 cd
2	484.4	77.3	850	20	Human SDCMP3 polyp
3	373.2	59.5	630	20	Mouse SDCMP3 polyp
4	373.2	59.5	1227	19	Mouse dectin-2 cdn
5	373.2	59.5	1252	22	Murine TANGO 405 a
6	344.8	55.0	821	22	Murine TANGO 405 c
7	303.2	48.4	501	19	Mouse dectin-2 ext
8	273.4	43.6	1312	22	Nucleotide sequenc
9	269.4	43.0	827	22	Dendritic cell (DC

10	265.6	42.4	800	22	AA019730	Dendritic cell (DC
11	207.2	33.0	324	22	AAK63162	Human immune/haema
12	201.2	32.1	444	24	ABK52914	Human cDNA encodin
13	200.4	32.0	402	24	ABK52903	Human cDNA encodin
14	192	30.6	1091	22	AAS31385	Human polynucleoti
15	192	30.6	1091	24	ABQ66709	Human cDNA encodin
16	192	30.6	1096	22	AAS31224	Human cDNA encodin
17	192	30.6	1096	24	ABQ66548	Human polynucleoti
18	192	30.6	1104	20	AA04885	Primate DCMPL C-le
19	192	30.6	1307	24	ABU90686	Human polynucleoti
20	176	28.1	582	24	ABK52904	Human dendritic ce
21	172.8	27.6	753	24	ABK52906	Human dendritic ce
22	166.8	26.6	1013	24	ABK52905	Human dendritic ce
23	151.4	24.1	558	24	ABK52907	Human dendritic ce
24	141.8	22.6	561	24	ABK52911	Human dendritic ce
25	135	21.5	448	24	ABK52912	Human dendritic ce
26	135	21.5	666	24	ABK52910	Human dendritic ce
27	134	21.4	568	24	ABK52908	Human dendritic ce
28	134	21.4	1418	20	AA04867	rodent DCMPL C-lec
29	130.2	20.8	403	24	ABK52909	Human dendritic ce
30	117.4	18.7	968	22	AA081743	Human membrane ass
31	117.4	18.7	997	20	AAK52274	Protein PRO244 CDN
32	117.4	18.7	997	22	AAK52936	Human DNA encoding
33	117.4	18.7	997	22	AA072432	Human PRO244 cDNA.
34	105.6	16.8	323	22	AAK53734	Murine transport a
35	96.6	15.4	559	24	ABK52913	Human dendritic ce
36	96	15.3	449	22	AAK63968	Human immune/haema
37	94.4	15.1	10409	19	AAV42558	Mouse dectin-2 gen
38	91	14.5	2517	22	AA086349	Murine cDNA encodi
39	77	12.3	2059	21	AA026386	Human secreted pro
40	60	9.6	2256	22	AA043054	Nucleotide sequenc
41	59.8	9.5	1521	23	AA071133	DNA encoding novel
42	59.8	9.5	2005	24	ABL95574	Human angiogenesis
43	59.8	9.5	2005	24	ABL88085	Human PRO223 cDNA
44	59.8	9.5	2024	21	AAA07697	Human collectin en
45	59.8	9.5	2181	22	AA160628	Human polynucleoti

ALIGNMENTS

RESULT 1  
AAS01375  
ID AAS01375 standard; cDNA; 3114 BP.

AC AAS01375;

XX 04-JUL-2001 (first entry)

DT Human TANGO 405 cDNA sequence.

DE Human; TANGO 210; clone jthLal52h06; TANGO 364; TANGO 366; dectin-2;

KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 154..783

FT /\*tag= a

FT /product= "TANGO 405 protein"

FT /note= "The ORF is specifically claimed"

FT sig\_peptide 154..297

FT /\*tag= b

FT mat\_peptide 298..780

FT /\*tag= c

PN WO200118016-A1.

XX 15-MAR-2001.

XX

PF 30-JUN-2000; 2000WO-US18174.  
XX PR  
XX 10-SEP-1999; 990S-0393996.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PR  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI: 2001-183280/18.  
DR P-PSDB; AAU00479.  
XX  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX  
XX Claim 2; Fig 6A-6C; 326pp; English.  
XX  
CC The present sequence encoding for human TANGO 405 protein is isolated  
CC from cDNA clone jthlal52h06 from a human mixed lymphocyte reaction cDNA  
CC library. It is 1 of 6 novel human proteins which include TANGO 210  
CC (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394  
CC (AAU00473), and INTERCEPT 400 (AAU00476). Novel sequences for murine  
CC TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and  
CC a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic  
CC acids encoding these novel proteins are useful as modulating agents in  
CC regulating a variety of cellular processes and can be used to express  
CC the proteins in a host cell in gene therapy applications. Human and  
CC murine TANGO 405 proteins show sequence homology to murine lectin-2.  
CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
CC activity, morphology and movement/migration of human lymphocytes and  
CC bone marrow cells and tissues and can be used to prevent, diagnose or  
CC treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 other;

Query Match 100.0%; Score 627; DB 22; Length 3114;  
Best Local Similarity 100.0%; Pred. No. 6.4e-193;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGCAAGCAGCAGCAACCTCAAAGTACAGAGAAAGAGCGCTGGTTGTCCTCGAGACTC 60  
DB 154 ATGATGCAAGCAGCAGCAACCTCAAAGTACAGAGAAAGAGCGCTGGTTGTCCTCGAGACTC 213  
QY 61 TGGTCTGTGGCTGGGATTCATTCAGTCACTCCTCAGTCTGCTTCAATGTGAGCTGTGTA 120  
DB 214 TGGTCTGTGGCTGGGATTCATTCAGTCACTCCTCAGTCTGCTTCAATGTGAGCTGTGTA 273  
QY 121 GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGCGCTGCTGAACCTACACTCATAT 180  
DB 274 GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGCGCTGCTGAACCTACACTCATAT 333  
QY 181 CATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA 240  
DB 334 CATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA 393  
QY 241 GCTTCTTGGAGTCAATTTGGTTCCAGTTGCTACTTCAATTTCCAGTGAAGAAGGTTTGG 300  
DB 394 GCTTCTTGGAGTCAATTTGGTTCCAGTTGCTACTTCAATTTCCAGTGAAGAAGGTTTGG 453  
QY 301 TCTAAGACTGACAGCAACTGCTTGACATGGGACACCATTTGCTGTGTTCAACACAGAA 360  
DB 454 TCTAAGACTGACAGCAACTGCTTGACATGGGACACCATTTGCTGTGTTCAACACAGAA 513  
QY 361 GCAGAGCAGAAATTTCAATGTCCAGCAGCTGAATGAGTCATTTTCTTATTTCTTGGGGCTT 420  
DB 514 GCAGAGCAGAAATTTCAATGTCCAGCAGCTGAATGAGTCATTTTCTTATTTCTTGGGGCTT 573  
QY 421 TCAGACCCACAGGTAATATAATGTCATGATTAAGACACCTTATGAGAAAAAT 480  
DB 574 TCAGACCCACAGGTAATATAATGTCATGATTAAGACACCTTATGAGAAAAAT 633  
QY 481 GTCAGATTTTGGCACCCTAGGTGAGCCCAATCATTCCTGACAGCAATGTGCTTCAATAGTC 540

DB 634 GTCAGATTTTGGCACCTAGTGAGCCCAATCATTTCTGCAGAGCAATGCTTCAATAGTC 693  
QY 541 TTCTGGAACCTACAGGATGGGCTGGGAATGATGTTATCTGTGAACTAGAGGAATTCA 600  
DB 694 TTCTGGAACCTACAGGATGGGCTGGGAATGATGTTATCTGTGAACTAGAGGAATTCA 753  
QY 601 ATATGTGAGATGAATAAGATTTCACCTA 627  
DB 754 ATATGTGAGATGAATAAGATTTCACCTA 780  
RESULT 2  
AAZ07531  
ID AAZ07531 standard; cDNA; 850 BP.  
XX  
XX AAZ07531;  
XX  
DT 26-NOV-1999 (first entry)  
XX  
DE Human SDCMP3 polypeptide encoding cDNA.  
XX  
KW Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;  
KW Schering dendritic cell membrane protein; dendritic cell physiology;  
KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
KW cancer; forensic; human; lectin 73; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 108..596  
FT /\*tag= a  
FT /product= "SDCMP3"  
XX  
PN WO9947673-A2.  
XX  
PD 23-SEP-1999.  
XX  
PF 16-MAR-1999; 99WO-US03740.  
XX  
PR 17-MAR-1998; 98US-0040111.  
XX  
PA (SCHE ) SCHERING CORP.  
PI Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
PI Philipps JH;  
XX  
DR WPI: 1999-562114/47.  
XX P-PSDB; AAY27446.  
XX  
PT Binding compound specific for primate or rodent Schering dendritic cell  
PT membrane proteins -  
XX  
PS Claim 8; Page 80; 89pp; English.  
XX  
CC The invention relates to a binding compound comprising an antibody  
CC binding site which specifically binds to: (a) a primate or rodent  
CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
CC primate SDCMP4 protein. The binding compound can be used to modulate  
CC dendritic cell physiology or function. The sequences may also be used as  
CC probes in forensic techniques, such as genetic fingerprinting. They can  
CC also be used to distinguish tissue and cell types in situ or in vitro.  
CC The sequences may also be useful in cancer immunotherapy and in the  
CC treatment of conditions associated with abnormal physiology or  
CC development, including abnormal proliferation, e.g. cancerous conditions,  
CC or degenerative conditions. The present sequence represents a cDNA  
CC encoding the human SDCMP3 (lectin 73).  
SQ Sequence 850 BP; 257 A; 153 C; 188 G; 252 T; 0 other;

Query Match 77.3%; Score 484.4; DB 20; Length 850;  
Best Local Similarity 98.8%; Pred. No. 7.5e-147;  
Matches 488; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAAGAGGCTGGTTGTCCTCCTGAGACTC	60
Db	108	ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAAGAGGCTGGTTGTCCTCCTGAGACTC	167
QY	61	TGGTCTGTGGCTGGGATTTCCATTGCACTCCTCAGTGCCTGCTTCATTGTGAGCTGTGTA	120
Db	168	TGGTCTGTGGCTGGGATTTCCATTGCACTCCTCAGTGCCTGCTTCATTGTGAGCTGTGTA	227
QY	121	GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGCTGAACACACATCAT	180
Db	228	GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGCTGAACACACATCAT	287
QY	181	CATTCAAGCTTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGCCCA	240
Db	288	CATTCAAGCTTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGCCCA	347
QY	241	GCTTCTTGGAGTCAATTTGGTTCAGTTGCTACTTCAATTTCCAGTGAAGAGCAAGGTTTG	300
Db	348	GCTTCTTGGAGTCAATTTGGTTCAGTTGCTACTTCAATTTCCAGTGAAGAGCAAGGTTTG	407
QY	301	TCTAAGAGTGAGCAGAACTGTGTTGAGATGGGACACATTTGTTGTTCACACAGAA	360
Db	408	TCTAAGAGTGAGCAGAACTGTGTTGAGATGGGACACATTTGTTGTTCACACAGAA	467
QY	361	GCAGAGCAGAAATTCATTTGTCAGCAGCTGAATGAGTCATTTTCTTATTTCTGGGCTT	420
Db	468	GCAGAGCAGAAATTCATTTGTCAGCAGCTGAATGAGTCATTTTCTTATTTCTGGGCTT	527
QY	421	TCAGACCCACAGGTAATAATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAT	480
Db	528	TCAGACCCACAGGTAATAATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAT	587
QY	481	GTCAGATTTTGGCA	494
Db	588	GTCAGGTGAGTGCA	601
RESULT 3			
ID	AAZ07532 standard; cDNA; 630 BP.		
XX	AAZ07532;		
AC	AAZ07532;		
DT	26-NOV-1999 (first entry)		
XX	Mouse SDCMP3 polypeptide encoding cDNA.		
DE			
XX			
KW	Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;		
KW	Schering dendritic cell membrane protein; dendritic cell physiology;		
KW	genetic fingerprinting; cancer immunotherapy; abnormal proliferation;		
KW	cancer; forensic; mouse; ss.		
XX			
OS	Mus sp.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..630	
FT	/*tag= a		
FT	/product= "SDCMP3"		
XX			
PN	W09947673-A2.		
XX			
XX	23-SEP-1999.		
PD			
XX	16-MAR-1999; 99WO-US03740.		
XX			
PR	17-MAR-1998; 98US-0040111.		
XX			
PA	(SCHE ) SCHERING CORP.		
XX			
PI	Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;		
PI	Phillips JH;		
XX			
DR	WPI; 1999-562114/47.		

DR	P-PSDB: AAY27447.		
XX	Binding compound specific for primate or rodent Schering dendritic cell		
PT	membrane proteins		
XX			
PS	Claim 8; Page 82; 89pp; English.		
XX			
CC	The invention relates to a binding compound comprising an antibody		
CC	binding site which specifically binds to: (a) a primate or rodent		
CC	Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a		
CC	primate SDCMP4 protein. The binding compound can be used to modulate		
CC	dendritic cell physiology or function. The sequences may also be used as		
CC	probes in forensic techniques, such as genetic fingerprinting. They can		
CC	also be used to distinguish tissue and cell types in situ or in vitro.		
CC	The sequences may also be useful in cancer immunotherapy and in the		
CC	treatment of conditions associated with abnormal physiology or		
CC	development, including abnormal proliferation, e.g. cancerous conditions,		
CC	or degenerative conditions. The present sequence represents a cDNA		
CC	encoding the mouse SDCMP3.		
XX			
SQ	Sequence 630 BP; 171 A; 135 C; 152 G; 172 T; 0 other;		
	Query Match	59.5%;	Score 373.2; DB 20; Length 630;
	Best Local Similarity	76.1%;	Pred. No. 9.2e-111;
	Matches 475; Conservative	0;	Mismatches 143; Indels 6; Gaps 1;
QY	10	GAGCAGCAACTCAAAAGTACAGAGAAAAGAGGCTGGTTGCTCGAGACTCTGGTCTGTG	69
Db	4	GTGCGAAGAAACAAATCCCAAGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCAGCT	63
QY	70	GCTGGGATTTCCATTGGACACTCCTCAGTGCCTGCTTCATTGTGAGCTGTAGTAACTTAC	129
Db	64	GCTGTGATTTCCATTGTTACTCTTGAGTACTGCTGTTTCATTGCGAGCTGTGGTGACTTAC	123
QY	130	CATTTTACATATGGTGAACACTGGCAAAAGGCTGTCTGAACACTACACTCATATCATTTCAAGT	189
Db	124	CAATTTATTATGGACCAGCCAGCTAGAAGACTATATGAACCTTCACACATACCATTCAGT	183
QY	190	CTCACCTGCTTCAGTGAAGGGACAAAGGTGCCAG-----CCTGGGGATGTTGCCAGCT	243
Db	184	CTCACCTGCTTCAGTGAAGGGACTATGGTGTGAGAAAAATGTGGGGATGCTGCCCAAT	243
QY	244	TCTTGAAGTCATTTGGTTCAGTTCCTACTTTCATTTCCAGTGAAGAGAGGTTTGGTCT	303
Db	244	CACCTGAAGTCATTTGGCTCCAGCTGCTACCTCATTTCTACCAAGAGAACTTCTGGAGC	303
QY	304	AGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTGTTCAACACAGAAGCA	363
Db	304	ACCAGTGAGCAGAACTGTGTTGAGATGGGGCTCATCTGGTGGTGATCAATCTGAAGCG	363
QY	364	GAGCAGAAATTCATTTGCCAGCAGCTGAATGAGTCATTTTCTTATTTCTGGGGCTTTCA	423
Db	364	GAGCAGAAATTCATCACCAGCAGCTGAATGAGTCATTTTCTTACTTCTGGGTCTTTGG	423
QY	424	GACCCACAAGGTAAATAAATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAATGTC	483
Db	424	GATCCACAAGGTAAATGGCAATGGCAATGGATCGATGATACCTCTTCAGTCAAAATGTC	483
QY	484	AGATTTTGGCACCTAGGTGAGCCCAATCATTTCTGAGAGCAATGTGCTCAATAGTCTTC	543
Db	484	AGGTTCTGGCACCCCAATGAACCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTTTAC	543
QY	544	TGGAACCTTACAGGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGGAATTCATAA	603
Db	544	TGGAATCTTCGAAATGGGGCTGGAATGATGTTTCTGTGATAGTAAACACAAATTCATA	603
QY	604	TGTGAGATGAATAGATTTACCTA	627
Db	604	TGTGAATGAAGAGATTTACCTA	627

RESULT 4

AAV42549

```

ID  AAV42549 standard; cDNA; 1227 BP.
XX
AC  AAV42549;
XX
DT  09-NOV-1998 (first entry)
XX
DE  Mouse dectin-2 cDNA.
XX
KW  Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;
KW  allergy; autoimmune disease; gene therapy; vaccine; diagnosis;
KW  drug screening; ss.
XX
OS  Mus sp.
XX
FH  Key Location/Qualifiers
FT  CDS 146..775
FT  /*tag= a
XX
XX  W09828332-A2.
XX
XX  02-JUL-1998.
XX
XX  22-DEC-1997; 97WO-US23761.
XX
XX  20-DEC-1996; 96US-0772440.
XX
XX  (TEXA ) UNIV TEXAS SYSTEM.
XX
XX  Arizumi K, Takashima A;
XX
XX  WP1; 1998-377594/32.
XX
XX  P-PSDB; AAW63010.
XX
XX  Nucleic acid encoding dendritic cell specific peptide(s) dectin-1
XX  and -2 - useful, e.g. to regulate immune response, as vaccine
XX  adjuvants, for diagnosis and drug screening
XX
XX  Claim 43; Page 141-142; 200pp; English.
XX
XX  This cDNA clone codes for a novel murine dendritic cell (DC)
XX  specific member of the C-type lectin family, termed DC-associated
XX  C-type lectin-2, or dectin-2 (see AAW63010), that is essential for
XX  DC-mediated T cell activation. Dectin-1 cDNA (see AAV42548) has
XX  also been identified. The invention provides: dectin-1 and
XX  dectin-2 polypeptides (see also AAW63009-22 and AAW69236-37), useful
XX  for purifying T cells, for detecting autoantibodies and for
XX  up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see
XX  AAV42548-54, AAV42558-60 and AAV44850-51); expression vectors;
XX  recombinant host cells; probes and primers useful e.g. for identifying
XX  human dectin-like molecules (see AAV42560); antibodies; compounds that
XX  modulate dectin-mediated activation of T cells; transgenic animals
XX  useful for studying dectin function and for drug screening; and
XX  dectin ligands. Dectin expression can be downregulated by
XX  antisense sequences or ribozymes, or by inactivating the dectin
XX  genes by homologous recombination, e.g. abrogation of dectin
XX  expression is useful for treating allergy and autoimmune
XX  disease. Alternatively gene disruption is done in vitro, the
XX  cells pulsed with an antigen, and returned to the patient to
XX  provide long-term non-responsiveness to the antigen. Since DC are
XX  'professional' antigen-presenting cells, genetic vaccines targeted
XX  to them should not induce tolerance.
XX
XX  Query Match 59.5%; Score 373.2; DB 19; Length 1227;
XX  Best Local Similarity 76.1%; Pred. No. 1.3e-110;
XX  Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;
XX
QY  10 GAGCAGCAACCTCAAGTACAGAGAAAGGCTGTGTCCCTGAGACTCTGGTCTGTG 69
DB  149 GTGCAGGAAAGACAATCCCAAGGAGGAGTCTGTGGACCCCTGAGACTCTGGTCAGCT 208
QY  70 GCTGGGATTCCTATGCACCTCCTCAGTGTGCTTTCATTCGTGAGCTGTGTAGTAACCTAC 129

```

PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI: 2001-183280/18.  
DR P-PSDB; AAU00482.

XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX Disclosure; Fig 6N-6P; 326pp; English.

XX The present sequence encoding for murine TANGO 405 alternative splice  
CC variant protein is isolated from cDNA clone jtmMa025all from a long-term  
CC bone marrow cDNA library. TANGO 405 (AAU00480) is 1 of 3 novel murine  
CC proteins which include TANGO 210 (AAU00470) and INTERCEPT 400  
CC (AAU00477). Six novel human proteins which include TANGO 210 (AAU00469),  
CC TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473),  
CC INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat  
CC INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids  
CC encoding these novel proteins are useful as modulating agents in  
CC regulating a variety of cellular processes and can be used to express  
CC the proteins in a host cell in gene therapy applications. Human and  
CC murine TANGO 405 proteins show sequence homology to murine dectin-2.  
CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
CC activity, morphology and movement/migration of human lymphocytes and  
CC bone marrow cells and tissues and can be used to prevent, diagnose or  
CC treat leukaemia, lymphomas and autoimmune disorders.

XX Sequence 1252 BP; 330 A; 283 C; 286 G; 353 T; 0 other;

Query Match 59.5%; Score 373.2; DB 22; Length 1252;  
Best Local Similarity 76.1%; Pred. No. 1.3e-110;  
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;

Qy 10 GAGCAGCAACCTCAAGTACAGAGAAAGAGCGTGGTGTCCCTGAGACTCTGGTCTGTG 69  
Db 182 GTGCAGGAAAGACAATCCCAAGGAGGAGTCTGTGGACCTGAGACTCTGGTCAAGCT 241  
Qy 70 GCTGGGATTTCCATGTCACCTCCTCAGTGTCTTCAATTCGAGCTGTAGTAACCTAC 129  
Db 242 GCTGTGATTTCCATGTTACTTCTGAGTACCTGTTCATTCGAGCTGTGGTCACTTAC 301  
Qy 130 CATTTTACATATGCTGAAATGGCAAAAGGCTGTCTGAACACACTCATATCATTCAGT 189  
Db 302 CAATTTATTTGGACCAAGCCAGTGAAGACTATATGAACCTTCACACATACCATTCAGT 361  
Qy 190 CTCACCTGCTTCAAGGAGCAAAAGTCCAG-----CCTGGGGATGTTGCCAGCT 243  
Db 362 CTCACCTGCTTCAAGGAGCAATGTTGTGTCAGAAAAAATGTGGGATGCTGCCCAAT 421  
Qy 244 TCTTGAAGTCATTTGGCTTCCAGTGTCTACTTTCATTTCCAGTGAAGAGGTTTGGTCT 303  
Db 422 CACTGGAAGTCATTTGGCTTCCAGTGTCTACTTCTATTTCTACCAAGGAGAACTTCTGGAGC 481  
Qy 304 AAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGTTTCAACAGCAAGCA 363  
Db 482 ACCAGTGAGCAGAACTGTGTTGAGATGGGAGCTCATCTGGTGTGATCATATCAAGCG 541  
Qy 364 GAGCAGAAATTTCAATGTCACAGCTGAATGAGTCATTTTCTATTTCTGGGGCTTTCA 423  
Db 542 GAGCAGAAATTTCAATCACCAGCAGCTGAATGAGTCACCTTTCTTACTTCTGGTCTTTCG 601  
Qy 424 GACCACAAGGTATATATATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAATGTC 483  
Db 602 GATCCACAAGGTATATGCAATGGCAATGGATGATGATATCTCTTTCAGTCAAAATGTC 661  
Qy 484 AGATTTTGGCACTTAGTGAGCCCAATCATCTCTGAGCAGCAATGTGCTTCAATAGTCTTC 543  
Db 662 AGGTTCTGGCAACCCCATGAACCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTTTAC 721  
Qy 544 TGAACACTCAGAGATGGGCTGGAATGATGTTATCTCTGTGAAACTAGAAGGAATTCATA 603  
Db 722 TGAATCTCTGAAATGGGCTGGAATGATGTTTCTCTGATAGTAAACACAAATTCATA 781

Qy 604 TCTGAGATGAATAAGATTTACCTA 627  
Db 782 TGTGAATGAAGAAGATTTACCTA 805  
RESULT 6  
ID AAS01376 standard; cDNA; 821 BP.  
XX AAS01376;  
AC AAS01376;  
XX 04-JUL-2001 (first entry)  
XX Murine TANGO 405 cDNA sequence.  
KW Murine; TANGO 210; clone jtmMa025all; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.  
XX Mus sp.  
XX Key Location/Qualifiers  
FH 174..710  
FT CDS /\*tag= a  
FT /\*product= "TANGO 405 protein"  
FT sig\_peptide 174..299  
FT /\*tag= b  
FT mat\_peptide 300..707  
FT /\*tag= c  
XX WO200118016-A1.  
XX 15-MAR-2001.  
XX 30-JUN-2000; 2000WO-US18174.  
XX 10-SEP-1999; 99US-0393996.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI: 2001-183280/18.  
XX P-PSDB; AAU00480.  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX Disclosure; Fig 6E-6F; 326pp; English.

XX The present sequence encoding for murine TANGO 405 protein is isolated  
CC from cDNA clone jtmMa025all from a long-term bone marrow cDNA library.  
CC An alternative splice variant of this protein is also described  
CC (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
CC INTERCEPT 400 (AAU00477). Six novel human proteins which include  
CC TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
CC INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
CC (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
CC described. The nucleic acids encoding these novel proteins are useful as  
CC modulating agents in regulating a variety of cellular processes and can  
CC be used to express the proteins in a host cell in gene therapy  
CC applications. Human and murine TANGO 405 proteins show sequence homology  
CC to murine dectin-2. TANGO 405 modulates growth, proliferation, survival,  
CC differentiation, activity, morphology and movement/migration of human  
CC lymphocytes and bone marrow cells and tissues and can be used to prevent,  
CC diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
XX Sequence 821 BP; 214 A; 188 C; 195 G; 223 T; 1 other;

PA	(TEXA ) UNIV TEXAS SYSTEM.		
XX			
PI	Ariizumi K, Takashima A;		
XX			
DR	WPI: 1998-377594/32.		
DR	P-PSDB; AAW63022.		
XX			
PT	Nucleic acid encoding dendritic cell specific peptide(s) dectin-1		
PT	and -2 - useful, e.g. to regulate immune response, as vaccine		
PT	adjuvants, for diagnosis and drug screening		
XX			
PS	Disclosure; Page 153; 200pp; English.		
XX			
CC	This cDNA sequence codes for the extracellular domain (see		
CC	AAW63022) of a novel murine dendritic cell (DC) specific member of		
CC	the C-type lectin family, termed DC-associated C-type lectin-2, or		
CC	dectin-2 (see also AAW63010), that is essential for DC-mediated T		
CC	cell activation. Full-length cDNA (see AAV42549) was isolated by		
CC	subtractive hybridisation from BALB/c mouse epidermis DC line XS52.		
CC	The invention provides: dectin-1 and -2 polypeptides (see also		
CC	AAW63009-22 and AAW69236-37), useful for purifying T cells, for		
CC	detecting autoantibodies and for up-regulating immunity e.g. as		
CC	vaccine adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and		
CC	AAV44850-51); expression vectors; recombinant host cells; probes		
CC	and primers; antibodies; compounds that modulate dectin-mediated		
CC	activation of T cells; transgenic animals; and dectin ligands.		
CC	Dectin expression can also be down-regulated to treat allergy and		
CC	autoimmune disease. Recombinant extracellular domain has been		
CC	expressed as a His-tagged protein (see AAW69237) useful for antibody		
CC	production.		
XX			
SQ	Sequence 501 BP; 142 A; 108 C; 115 G; 136 T; 0 other;		
Query Match 48.4%; Score 303.2; DB 19; Length 501;			
Best Local Similarity 78.4%; Pred. No. 4.7e-88;			
Matches 378; Conservative 0; Mismatches 98; Indels 6; Gaps 1;			
QY	152	GCAAAAGGCTGTGTAACCTACACTCATATCATTAAGTCTCACCTGCTTCAGTGAAGGGA 211	
DB	20	GTAGAAGACTATATGAACCTTCACATACATACCTCCAGTCTCACCTGCTTCAGTGAAGGGA 79	
QY	212	CARAAGTGGCAG-----CCTGGGGATGTGCCAGCTTCTTGGAGTCATTTGGTTCCA 265	
DB	80	CTATGTGTGCAGAAAAAATGTGGGGATGCTGCCCAAACTCACTGGAAGTCAATTTGGCTCCA 139	
QY	266	GTTGCTACTTTCATTTCCAGTGAAGAGAAGGTTTGGTCTTAAGAGTCAGCAGAACTGTGTTG 325	
DB	140	GCTGCTACCTCATTTCTACCAAGGAGAACTTCTGGAGCACCACTGAGCAGAACTGTGTTT 199	
QY	326	AGATGGGAGCACAATTTGGTTGGTTTCAACACAGACAGCAGAGAAATTTCAATTTGCCAGC 385	
DB	200	AGATGGGGCTCATCTGTGGTGTATCAATACTGAAGCGGAGAGAAATTTTCATCACCCAGC 259	
QY	386	AGCTGAATGAGTCATTTCTTATTTCTGGGGCTTTTCAGACCCACACAAGCTAATAATAATT 445	
DB	260	AGCTGAATGAGTCATTTCTTACTTCTCTGGTCTTTTCGGATCCACAAGAGTGAATGGCAAT 319	
QY	446	GGCAATGGATTGATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGC 505	
DB	320	GGCAATGGATCGATGATACTCTTTTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAAC 379	
QY	506	CCAATCATTTCTGCAGACAAATGTGCTTCAATAGTCTTCTGGAAACCTCAGAGATGGGGCT 565	
DB	380	CCAATCTTCCAGAAGACGGTGTGTTTCAATAGTCTTCTGGAATCTCTGGAATGGGGCT 439	
QY	566	GAATGATGTTATCTGTGAAACTAGAAGGAATTCATATGTGAGATGAATAAGATTTACC 625	
DB	440	GAATGATGTTTCTGTGATAGTAAACACAATTCATTAATGTGAATGAAGAATTTACC 499	
QY	626	TA 627	
DB	500	TA 501	

RESULT 8  
AAF90241  
ID AAF90241 standard; cDNA; 1312 BP.  
XX  
AC AAF90241;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Nucleotide sequence of BDCA-2 antigen.  
XX  
KW BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;  
KW viral infection; autoimmune disease; allergic response; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 491..1132  
FT FT /\*tag= a  
FT FT /product= "BDCA-2"  
XX  
XX WO200136487-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 15-NOV-2000; 2000WO-1B01832.  
XX  
PR 15-NOV-1999; 99US-0165555.  
PR 23-NOV-1999; 99US-0167076.  
PR 28-JAN-2000; 2000US-0179003.  
PR 07-FEB-2000; 2000US-0180775.  
PR 11-APR-2000; 2000US-0196824.  
PR 13-APR-2000; 2000US-0197205.  
XX  
XX (MILT-) MILTENYI BIOTECH GMBH.  
XX  
XX Schmitz J, Dzionek A, Buck DW;  
XX  
XX WPI: 2001-355622/37.  
XX P-PSDB; AAB84215.  
XX  
XX Compositions and cell populations enriched in dendritic cells through  
XX use of antigen-binding fragments specific for BDCA-2, BDCA-3 or  
XX BDCA-4, are used to treat viral infections, autoimmune diseases,  
XX allergic responses and cancer -  
XX  
XX Claim 134; Fig 12; 115pp; English.  
XX  
XX The present sequence encodes an antigen designated BDCA-2. The  
XX specification describes compositions and hematopoietic cell populations  
XX enriched in dendritic cells (DCs). These compositions are produced using  
XX antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4. The DCs  
XX obtained are used to treat viral infections, autoimmune disease, allergic  
XX response, and cancer. BDCA-1, BDCA-2, BDCA-3 and BDCA-4 monoclonal  
XX antibodies and their antigen-binding fragments are used to detect,  
XX enumerate and isolate DC populations from leukapheresis material, whole  
XX blood and tonsils and from non-hematopoietic and hematopoietic tissues.  
XX  
XX Sequence 1312 BP; 374 A; 301 C; 297 G; 340 T; 0 other;  
XX

Query Match 43.6%; Score 273.4; DB 22; Length 1312;  
Best Local Similarity 66.7%; Pred. No. 3.8e-78;  
Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;  
Qy 1 ATGATGCAAGACGACCAACCTCAAGTACAGAGAAAGAGCGGTGGTGTCC---TGAGA 57  
Dy 491 ATGGTGCTGAAGAAGAGCCTCAAGACCGAGAGAAAGAGCTCTGGTGGTTCCAGTTGAAG 550  
Qy 58 CTCTGGCTCTGGCTGGGATTTCCATTGCACCTCTCAGTGCCTTCATTGTCAGCTGT 117  
Dy 551 GTCTGGTCCATGGCGAGTCGATCCATCTTGTCTCTCAGTGTCTGTCTTCTGAGTTCT 610  
Qy 118 GTAGTAACCTTACCATTATATATGTTGTAAGTGGCAAAAGGCTGTCTGAACCTA----- 171

Db 611 GTGGTGCCTCAACAATTTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAGTTACGAGAG 670  
Qy 172 ---CACTCATATCATTCAGTCTCAACCTGCTTCAAGTGAAGGACAAAGGTGCCAGCCTGG 228  
Db 671 TATCAACAGTATCATCCAAGCCTGACCTGGCTCAAGTGAAGGACAAAGACATAGAAGATTGG 730  
Qy 229 GGATGTTGCCAGCTTCTTTGGAAGTCATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAA 288  
Db 731 AGCTGCTGCCCAACCCCTTGGACTTCATTTCACTCTAGTTGCTACTTATTTCTACTGGG 790  
Qy 289 GAGAAGGTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGCAGCATTTGGTTGTG 348  
Db 791 ATGCAATCTTGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGGCTGATCTGGTGGTG 850  
Qy 349 TTCAACACAGAAGCAGACAGAAATTTCAATTTGTCACAGAGCTCAATGAGTCATTTTCTTAT 408  
Db 851 ATCAACACAGGGAAGACAGGATTTTCATTCATTCAGAACTCTGAAAAGAAATTTCTTCTTAT 910  
Qy 409 TTTCTGGGGCTTTTCAGAGCCCAAGGTAATAATAATTTGCAATGATGATGATAAGACACT 468  
Db 911 TTTCTGGGGCTGTCAAGTCCAGGGGCTCGGCACATTGGCAATGGTTTGACCAGACACCA 970  
Qy 469 TATGAGAAAATGTCAGATTTTGGCACCTAGTGTAGTGAGGCCAAATCATTTCTGCAGAGCAATGT 528  
Db 971 TACAATGAAAATGTCACATTTCTGGCACCTCAGGTGAACCAATAACCTTGTATGAGCGGTTGT 1030  
Qy 529 GCTTCAATAGTCTTCTGGAAACCTACAGATGGGGCTGGAATGATGTTTATCTGTCAAACT 588  
Db 1031 GCGATAATAATTTCCGTTCTTTCAAGAAATGGGGCTGGAATGACATTTCTGTCATGTA 1090  
Qy 589 AGAAGGAATTCATATGTGAGATGAATAAGATTTTACCTA 627  
Db 1091 CCTCAGAGTCAATTTGCAAGATGAAGAAGATCTACATA 1129

RESULT 9  
AAD19729  
ID AAD19729 standard; DNA; 827 BP.  
XX  
AC AAD19729;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Dendritic cell (DC) DCLEC gene.  
XX  
KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
KW transplant rejection; chronic inflammatory disease; anti-HIV; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..657  
FT FT /\*tag= a  
FT FT /product= "Dendritic cell DCLEC protein"  
XX  
XX WO200127773-A2.  
XX  
XX 04-OCT-2001.  
PD  
PF 28-MAR-2001; 2001WO-EP03542.  
XX  
XX 29-MAR-2000; 2000US-192934P.  
PR 18-MAY-2000; 2000US-205020P.  
PR 18-MAY-2000; 2000US-205026P.  
PR 19-MAY-2000; 2000US-205767P.  
PR 19-MAY-2000; 2000US-205769P.  
XX  
XX (NOVS ) NOVARTIS AG.  
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX PI Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
XX DR WPI; 2001-616466/71.  
XX DR P-PSDB; AAE12079.  
XX PT New polypeptides for screening therapeutic agonists and antagonists  
XX PT comprise dendritic cell polypeptides -  
XX PS Claim 1; Page 50-51; 52pp; English.  
XX CC The invention relates to dendritic cell (DC) proteins and their  
XX CC corresponding DNA molecules. A pharmaceutical composition comprising  
XX CC agonist and antagonist of DC proteins are useful for treating abnormal  
XX CC conditions related to both an excess of and insufficient level of  
XX CC expression of DC gene, or related to both an excess of and insufficient  
XX CC activity of DC protein. Soluble form of DC proteins are used as an active  
XX CC ingredient in combination with pharmaceutical acceptable carriers.  
XX CC DC genes and proteins are useful for treating chronic inflammatory  
XX CC diseases, autoimmune diseases, transplant rejection crisis, including  
XX CC inflammatory skin diseases such as contact hypersensitivity, atopic  
XX CC dermatitis or virally-induced immune suppression such as AIDS and cancer.  
XX CC DC protein is useful for inducing immunological response in a mammal, and  
XX CC as immunogen to produce antibodies immunospecific for the polypeptide.  
XX CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
XX CC reagent, and for chromosomal identification. The present sequence is  
XX CC dendritic cell (DC) DCLEC gene which is found to belong to the family  
XX CC of C-type lectins with one single carbohydrate recognition domain at the  
XX CC C-terminal end.  
SQ Sequence 827 BP; 242 A; 164 C; 188 G; 233 T; 0 other;

Query Match 43.0%; Score 269.4; DB 22; Length 827;  
Best Local Similarity 66.8%; Pred. No. 6e-77;  
Matches 419; Conservative 0; Mismatches 196; Indels 12; Gaps 2;

QY 13 CAGCAACCTCAAGTACAGAAAGAGGCTGGTGTGCC---TGAGACTCTGGTGTGG 69  
DB 28 CAAGAGCCTCAAGACCGAGAGAGGAGCTGGTGGTTCCAGGTTGAAGGCTGGTCCATG 87  
QY 70 GCTGGGATTTCCATTGCTCACTCAGTCTGCTTCATTGTGAGTGTAGTAACCTTAC 129  
DB 88 CGAGTCTGATCCATCTTCTCCTCAGTCTGCTTTCCTCAGTGTGAGTCTGTGGTGCCTCAC 147  
QY 130 CATTTTACATATGGTGAACCTGGCAAGAGGCTGTCTGAACCTA-----CACTCATAT 180  
DB 148 AATTTTATGATAGCAAAACTGTCAAGAGGCTGTCCAAAGTTACGAGAGTATCAACAGTAT 207  
QY 181 CATTCAGCTCACCTGCTTCAGTGAAGGGACAAAGGTGCCAGCTGGGGATGTGCCCA 240  
DB 208 CATCCAAGCCTGACCTGGCTCATGGGAAGGACATAGAGATTTGGAGCTGCTGCCCA 267  
QY 241 GCTTCTTGGAGTCAATTTGGTTCCAGTTGCTACTTTCATTTCAGTGAAGAGAGGTTGG 300  
DB 268 ACCCTTGGACTTCATTTTCAGTCTAGTGTCTACTTTATTTCTACTGGGATCAATCTTGG 327  
QY 301 TCTAAGATGAGCAGACTGTGTTGAGATGGGAGCACATTTGGTTGTGTTCAACACAGAA 360  
DB 328 ACTAAGAGTCAAAAGAACTGTTCTGTATGGGGGTGATCTGGTGGTATCAACACCAGG 387  
QY 361 GCAGAGCAGAATTTTCATTTGCCAGCAGTGAATGAGTCATTTTCTATTTTCTGGGCTT 420  
DB 388 GAAGAACAGGATTTTCATTCATTCAGTAATCTGAAAGAAATTTCTTTATTTTCTGGGCTG 447  
QY 421 TCAGACCCACAGGTAATAAATATGGCAATGGATTGATTAAGACACACCTTATGAGAAAAT 480  
DB 448 TCAGATCCAGGGGTGCGGACATTTGGCAATGGTTGACACAGACACCAATGAAAAT 507  
QY 481 GTCATTTTGGCAGCTAGGTGAGCCCCAATCATTTCTGACAGCAATGTGCTTCAATAGTC 540  
DB 508 GTCATTTTGGCAGCTAGGTGAGCCCCAATCAACCTTGTATGATGCGGTGTGCGCAATAAAT 567  
QY 541 TTCTGGAACCTACAGGATGGGGCTGGAAATGATGTTATCTGTGAACAGTAAGAAATCA 600

Db 568 TTCGGTTCTTCAGAGAATGGGCTGGAATGACATTCACCTGTCATGTCACCTCAGAAGTCA 627  
QY 601 ATATGTGAGATGAATAAGATTTACCTA 627  
Db 628 ATTGCAAGATGAAGAAGATCTACATA 654  
RESULT 10  
AAD19730  
ID AAD19730 standard; DNA; 800 BP.  
XX AC AAD19730;  
XX DT 18-DEC-2001 (first entry)  
XX DE Dendritic cell (DC) DCLEC/SPLICE 1 variant gene.  
XX KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
XX KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
XX KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
XX KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
XX KW transplant rejection; chronic inflammatory disease; anti-HIV; variant;  
XX KW ds.  
XX OS Unidentified.  
XX FH Key  
XX FT CDS  
XX FT Location/Qualifiers  
XX FT 1..634  
XX FT /\*tag= a  
XX FT /product= "dendritic cell DCLEC/SPLICE 1 variant protein"  
XX FT /transl\_except= (pos:1, aa:Pro)  
XX FT /note= "This codon has an apparent deletion of 2  
XX FT nucleotides, which alters the reading frame; CDS does not  
XX FT include start codon"  
XX FT /partial  
XX FT  
XX WO200172773-A2.  
XX PD 04-OCT-2001.  
XX PF 28-MAR-2001; 2001WO-BP03542.  
XX PR 29-MAR-2000; 2000US-192934P.  
XX PR 18-MAY-2000; 2000US-205020P.  
XX PR 18-MAY-2000; 2000US-205026P.  
XX PR 19-MAY-2000; 2000US-205767P.  
XX PR 19-MAY-2000; 2000US-205769P.  
XX (NOVS ) NOVARTIS AG.  
XX PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
XX PI Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
XX WPI; 2001-616466/71.  
XX P-PSDB; AAE12080.  
XX New polypeptides for screening therapeutic agonists and antagonists  
XX comprise dendritic cell polypeptides -  
XX Claim 1; Page 51; 52pp; English.  
XX CC The invention relates to dendritic cell (DC) proteins and their  
XX CC corresponding DNA molecules. A pharmaceutical composition comprising  
XX CC agonist and antagonist of DC proteins are useful for treating abnormal  
XX CC conditions related to both an excess of and insufficient level of  
XX CC expression of DC gene, or related to both an excess of and insufficient  
XX CC activity of DC protein. Soluble form of DC proteins are used as an active  
XX CC ingredient in combination with pharmaceutical acceptable carriers.  
XX CC DC genes and proteins are useful for treating chronic inflammatory  
XX CC diseases, autoimmune diseases, transplant rejection crisis, including  
XX CC inflammatory skin diseases such as contact hypersensitivity, atopic  
XX CC dermatitis or virally-induced immune suppression such as AIDS and cancer.



CC DC protein is useful for inducing immunological response in a mammal, and  
CC as immunogen to produce antibodies immunospecific for the polypeptide.  
CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
CC reagent, and for chromosomal identification. The present sequence is  
CC dendritic cell (DC) DCLEC/SPICE-1 variant gene which is found to belong  
CC to the family of C-type lectins with one single carbohydrate recognition  
CC domain at the C-terminal end.  
XX

SQ Sequence 800 BP; 231 A; 160 C; 183 G; 226 T; 0 other;

Query Match 42.4%; Score 265.6; DB 22; Length 800;  
Best Local Similarity 66.8%; Pred. No. 1e-75; Indels 9; Gaps 1;  
Matches 398; Conservative 0; Mismatches 189;

QY 41 GCTGTTGCTCCCTGAGACTCTGTCTCTGTGGCTGGGATTTCCATTTGCATTCCTCAGTGGCT 100  
DB 36 GGTGGTTCACGTTCAAGGCTGTGGTCCATGTCATCCATCTTGTCTCCTCAGTGTCT 95  
QY 101 GCTTCATTTGAGCTGTGTAGTAACTTACCATTTTACATATGTTGAACTGGCAAGGC 160  
DB 96 GTTTCACGTGAGTGTGTGGTGGTCCACAAATTTATGATAGCAAACTGTCAAGAGGC 155  
QY 161 TGTCTGAACTA-----CACTCATATCATTTCAAGTCTCACCTGCTTTCAGTGAAGGA 211  
DB 156 TGTCCAGTTACGAGAGATATCAACAGTATCATCCAAAGCTGACCTGCGTCATGGAAGAA 215  
QY 212 CAAGGTCCAGCTGGGGATGTTGCCAGCTTCTTGGAGTCAATTTGGTTCCAGTTGCT 271  
DB 216 AGGACATAGAGATGTTGGAGTGTGCCCCAACCCCTTGGACTTCATTTTCAGTCTAGTTGCT 275  
QY 272 ACTTCATTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTCAGCAGAACTGTGTTGAGATGG 331  
DB 276 ACTTTATTTCTACTGGGATGCAATCTTGGACTTAAGAGTCAAAAGAACTCTTCTGTGATGG 335  
QY 332 GAGCACATTTGGTGTGTTGCACACAGACGACGAGAAATTTCAATTTGCCAGAGCTGA 391  
DB 336 GGGCTGTATGTTGGTGTGATCAACACACGAGGAGCAAGAGATTTCAATTTCAAGATCTGA 395  
QY 392 ATGAGTCAATTTCTATTTCTGGGGCTTTCAGACCCACCAAGGTAATATATTTGGCAAT 451  
DB 396 AAGAATAATCTTCTATTTCTGGGGCTGTACATCCAGGGGTGCGGCACATTTGGCAAT 455  
QY 452 GGATTTGATAGACACCTTATGAGAAATGTCAGAAATTTTGGACCTAGTGAGGCCAATC 511  
DB 456 GGGTTGACGACACACCACTACAATGAAATGTCACATTTCTGGCACTCAGTGAACCAATA 515  
QY 512 ATCTGACAGCAATGTGCTTCAATAGTCTTCTGGAACCTACAGCATGGGCTGGAATG 571  
DB 516 ACCTTGATGAGGTTGTGGATATAAATTTCCGTTCTTTCAGAGAAATGGGCTGGAATG 575  
QY 572 ATGTTATCTGTAACCTAGAGAGAAATTCATATGTGAGATGAATAGATTTACCTA 627  
DB 576 ACATTTCACTGTCATGTACCTCAGAGTCAATTTGCAAGATGAAGAGATCTACATA 631

RESULT 11

AAK63162

ID AAK63162 standard; cDNA; 324 BP.

XX AC AAK63162;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8222.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX OS cytostatic; gene therapy; vaccine; metastasis; ss.

XX PN Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216680.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0232399.  
XX 14-SEP-2000; 2000US-0232400.  
XX 14-SEP-2000; 2000US-0232401.  
XX 14-SEP-2000; 2000US-0233063.  
XX 14-SEP-2000; 2000US-0233064.  
XX 14-SEP-2000; 2000US-0233065.  
XX 21-SEP-2000; 2000US-0234223.  
XX 21-SEP-2000; 2000US-0234274.  
XX 25-SEP-2000; 2000US-0234997.  
XX 25-SEP-2000; 2000US-0234998.  
XX 26-SEP-2000; 2000US-0235484.  
XX 27-SEP-2000; 2000US-0235834.  
XX 27-SEP-2000; 2000US-0235836.  
XX 29-SEP-2000; 2000US-0236327.  
XX 29-SEP-2000; 2000US-0236367.  
XX 29-SEP-2000; 2000US-0236368.

```
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-483426/52.
XX P-PSDB; RAM90381.
XX

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Claim 1; SEQ ID NO 8222; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK94950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 324 BP; 84 A; 72 C; 80 G; 86 T; 2 other;
Query Match 33.0%; Score 207.2; DB 22; Length 324;
Best Local Similarity 91.3%; Pred. No. 6.1e-57;
Matches 251; Conservative 1; Mismatches 20; Indels 3; Gaps 3;
QY 1 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGTGGTTGTCCTGAGACTC 60
Db |||||
QY 49 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGTGGTTGTCCTGAGACTC 108
Db |||||
QY 61 TGGTCTGTGGCTGGGATTCATTCGCACTCTCAGTCGCTTG-CTTCATTGTCAGCTGTGT 119
Db |||||
QY 109 TGGTCTGTGGCTGGGATTCATTCGCACTCTCAGTCGCTTGCTGATTCATTCAGCTGTGT 168
Db |||||
QY 120 AGTAACCTTACCATTTCATATGCTGAACTGG-CAAAAGCGTCTCTGAACACACATCAT 178
Db |||||
QY 169 AGTAACCTTACCATTTCATATGCTGAACTGG-CAAAAGCGTCTCTGAACACACATCAT 228
Db |||||
QY 179 ATCATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTCAGCCT-GGGGATGTTGC 237
Db |||||
QY 229 ATCATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTCAGCCTGGGGATGTTGC 288
Db |||||
QY 238 CCAGCTTCTTGAAGTCAATTTGGTTCCAGTTGCTA 272
Db |||||
QY 289 CCAGCTTCTTGAAGTCAATTTGGTTCCAGTTGCTA 323
Db |||||
RESULT 12
ABK52914
ID ABK52914 standard; cDNA; 444 BP.
XX
XX ABK52914;
XX
XX 27-AUG-2002 (first entry)
XX
XX Human cDNA encoding a partial dendritic cell immunoreceptor #2.
XX
XX Human; ss; gene; dendritic cell immunoreceptor; cytostatic;
XX antiasthmatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;
XX asthma; inflammation; obesity; diabetes; central nervous system disorder;
XX Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;
XX cardiovascular disorder; myocardial infarction; ischaemic heart disease;
XX congestive heart failure; chronic obstructive pulmonary disease; COPD.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..444
XX
```



CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
CC disorders are listed in the specification). The polypeptide and nucleic  
CC acid are useful for identifying test compounds which act as agonists or  
CC antagonists, for raising specific antibodies, and as a bait protein in a  
CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
CC diagnostic assays for detecting diseases and abnormalities or  
CC susceptibility to disease and abnormalities related to the presence of  
CC mutations. The present sequence is a cDNA encoding a partial  
CC dendritic cell immunoreceptor.

XX  
SQ Sequence 402 BP; 114 A; 80 C; 96 G; 112 T; 0 other;

Query Match 32.0%; Score 200.4; DB 24; Length 402;

Best Local Similarity 68.7%; Pred. No. 1.1e-54;

Matches 276; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 226 TGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCAGTTGCTACTTCAATTTCCAGT 285

Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTGCTACTTTATTTCTACT 60

QY 286 GAAGAGAAGTTTGGTCTTAAGAGTGAGCAGAACTGTGTGAGATGGAGCACATTTGGTT 345

Db 61 GGGATCAATCTTGGACTAAGAGTCAAAAGAACTGTCTGTGATGGGGCTGATCTGGTG 120

QY 346 GTCTTCAACACAGACAGACAGAAATTTTCATTTGCCAGCAGCTGAATGAGTCATTTTCT 405

Db 121 GTATCAACACAGGAGAAACAGGATTTTCATTCAGAAATCTGAAAGAAATTTCTTCT 180

QY 406 TATTTTCTGGGGCTTTTCAGACCCACAGGTAATAATATTTGGAAATGGATTTGATGACA 465

Db 181 TATTTTCTGGGGCTGTTCAGATCCAGGGGTTCGGCGACATTTGCAATGGTTGACCAGACA 240

QY 466 CTTATGAGAAATGTCAGATTTTGGCACTAGGTGAGCCCAATCATTTCTGCAGAGCAA 525

Db 241 CCATCAATGAAATGTCAGATTTTGGCACTAGGTGAGCCCAATCATTTCTGCAGAGCAA 300

QY 526 TGTGCTTCAATAGTCTTCTTGGAAACCTACAGGATGGGGTGAATGATGTTATCTGTGAA 585

Db 301 TGTGGGATAATAATTTCCGTTCTTCAGAAGAATGGGCTGGAATGACATTCACCTGTCAT 360

QY 586 ACTAGAAGGAATCAATATGTGAGATGAATAAGATTTACCTA 627

Db 361 GTACCTCAGAAGTCAATTTTCAAGATGAAGAAGATCTACATA 402

RESULT 14

AAS31385

ID AAS31385 standard; cDNA; 1091 BP.

XX

AC AAS31385;

XX

DT 04-DEC-2001 (first entry)

XX

DE Human cDNA encoding a novel extracellular matrix protein, Seq ID No 199.

XX

KW Human; secreted extracellular matrix protein; ss; immunomodulatory;  
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.

XX

OS Homo sapiens.

XX

PN WO200155368-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01348.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 21-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 17, 2003, 08:49:16 : Search time 22.9615 Seconds  
(without alignments)  
13907.882 Million cell updates/sec

Title: US-09-766-511B-52  
Perfect score: 627  
Sequence: 1 atgatgaagagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues  
Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	355.6	56.7	693	10	US-09-833-381-994
c 2	311.6	49.7	528	10	US-09-833-381-995
c 3	273.4	43.6	642	12	US-10-090-466-1
4	231	36.8	549	12	US-10-090-466-3
5	192	30.6	1091	10	US-09-764-870-199
6	192	30.6	1096	10	US-09-764-870-38
7	192	30.6	1104	9	US-09-862-802-1
c 8	181.8	29.0	1036	10	US-09-833-381-119
c 9	151.2	24.1	291	10	US-09-833-381-996
10	134	21.4	1418	9	US-09-862-802-7
11	120.6	19.2	758	10	US-09-833-381-1339
12	117.4	18.7	968	9	US-09-965-529-40
13	117.4	18.7	997	9	US-09-905-291A-376
14	117.4	18.7	997	9	US-09-902-853-376
15	117.4	18.7	997	9	US-09-907-824-376
16	117.4	18.7	997	9	US-09-907-841-376
17	117.4	18.7	997	9	US-09-904-011-376
18	117.4	18.7	997	9	US-10-174-590-23
19	117.4	18.7	997	9	US-10-176-758-23

20	117.4	18.7	997	9	US-10-175-737-23	Sequence 23, Appl
21	117.4	18.7	997	9	US-09-906-742-376	Sequence 376, App
22	117.4	18.7	997	9	US-10-173-706-23	Sequence 23, Appl
23	117.4	18.7	997	9	US-10-175-738-23	Sequence 23, Appl
24	117.4	18.7	997	9	US-10-175-752-23	Sequence 23, Appl
25	117.4	18.7	997	9	US-10-176-482-23	Sequence 23, Appl
26	117.4	18.7	997	9	US-10-176-757-23	Sequence 23, Appl
27	117.4	18.7	997	9	US-10-176-913-23	Sequence 23, Appl
28	117.4	18.7	997	9	US-10-180-552-23	Sequence 23, Appl
29	117.4	18.7	997	9	US-10-180-557-23	Sequence 23, Appl
30	117.4	18.7	997	9	US-09-906-838-376	Sequence 376, App
31	117.4	18.7	997	9	US-09-907-613-376	Sequence 376, App
32	117.4	18.7	997	9	US-09-907-942-376	Sequence 376, App
33	117.4	18.7	997	9	US-10-173-700-23	Sequence 23, Appl
34	117.4	18.7	997	9	US-10-174-572-23	Sequence 23, Appl
35	117.4	18.7	997	9	US-10-174-579-23	Sequence 23, Appl
36	117.4	18.7	997	9	US-10-174-582-23	Sequence 23, Appl
37	117.4	18.7	997	9	US-10-174-588-23	Sequence 23, Appl
38	117.4	18.7	997	9	US-10-175-739-23	Sequence 23, Appl
39	117.4	18.7	997	9	US-10-175-740-23	Sequence 23, Appl
40	117.4	18.7	997	9	US-10-175-743-23	Sequence 23, Appl
41	117.4	18.7	997	9	US-10-176-488-23	Sequence 23, Appl
42	117.4	18.7	997	9	US-10-176-492-23	Sequence 23, Appl
43	117.4	18.7	997	9	US-10-176-747-23	Sequence 23, Appl
44	117.4	18.7	997	9	US-10-176-750-23	Sequence 23, Appl
45	117.4	18.7	997	9	US-10-176-985-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-09-833-381-994/c  
; Sequence 994, Application US/09833381  
; Patent No. US2002132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US2002132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 994  
; LENGTH: 693  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(693)  
; OTHER INFORMATION: n = A,T,C or G  
; US-09-833-381-994

Query Match	56.7%	Score 355.6;	DB 10;	Length 693;
Best Local Similarity	76.5%;	Pred. No. 1.6e-99;		
Matches 463;	Conservative 0;	Mismatches 135;	Indels 7;	Gaps 2;
QY	30	AGAGAAAGAGGCTGCTTCCTCCCTGAGACTCTGGTCTGGTGGGATTTCACATTGCACCT 89		
Db	690	AGGAGAGGAGTCTGCTGGACCTGAGACTCTGGTCTGGTCTGGTCTGCTGATTCACATTGTTACT 631		
QY	90	CCTCAGTGTCTGCTTCATTGTGAGCTGTGTAGTAACCTTACCATTTCACATATGTTGGAAC 149		
Db	630	CTTGACTACCTGTTTCATTTCGCGAGCTGTGGTGGTACTTACCAATTTATTTATGGACGACC 571		
QY	150	TGCAAAAGGCTCTCTGAAGTACACTCATATCATTCACAGTCTCACCTGCTTCAGTGAAGG 209		
Db	570	CAGTAGAAGACTATATGAACCTTCACACATACCATTCACAGTCTCACCTGCTTCAGTGAAGG 511		
QY	210	GACAAAGTGGCAG-----CCTGGGGATGTTGGCCAGCTTCTTGAAGTCAATTGGTTC 263		

[illegible]







```
Db 371 GGTGGCGGACATTTGGCAATGGTTTGACGAGACACCACATACATGAAATGTCCAGTGGAGTA 312
Qy 484 -----AGATTTTGGCACCTAGGTGAGCCCAATCTTCGACAGCAATGTGCTTCAAT 536
Db 311 TAGAATGAGATTCGGCACTAGCTGAGCAACCAATAACCTTGATGAGCGGTGTGGCGATAAT 252
Qy 537 AGTCTTCTGGAAACCTACAGGATGGGCTGGAATGATGTTATCTGTGTAACACTAGAAGGAA 596
Db 251 AAATTTCCGTTCTTCAGAAGAATGGGCTGGAATGACATTCACCTGTCATCTACCTCAGAA 192
Qy 597 TTCAATATCTGAGATGAATAGATTTACCTA 627
Db 191 GTCAATTTGCAAGATGAAGAGATCTACATA 161

RESULT 9
US-09-833-381-996/c
; Sequence 996, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 996
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(291)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-996

Query Match 24.1%; Score 151.2; DB 10; Length 291;
Best Local Similarity 72.5%; Pred. No. 7.5e-37;
Matches 211; Conservative 0; Mismatches 78; Indels 2; Gaps 2;

Qy 264 CAGTGTCTACTTTCATTTCCAGTGAAG-AGAAAGTTTGGTCTAAGAGTGAGCAGCAACTGTG 322
Db 291 CAGCTGTACTCTCATTTCTACCAGGAAGAACITCTCNAGCACCAGTGAGCANAACGTNG 232
Qy 323 TTGAGATGGGAGCACATTTGGTTGTGTCA-ACACAGAAGCAGACGAGAAATTTTCATTTGC 381
Db 231 TTCATATGGGGGCTCATCTGGNGGTGATCACATACTCAAGCGGAGCAGANTTTTCATCACC 172
Qy 382 CAGCAGCTGAATGAGTCATTTCTTATTTCTGGGGCTTTCAGACCCACAAAGGTAATAAT 441
Db 171 CAGCAGTGAATGAGTCACITTTTACTTCTCGGTCTTTTCGGATCCACAAAGGTNNNGG 112
Qy 442 AATTGGCAATGGATTGATAAGACACCTTATGAGAAATATGTCAGATTTTGGCACCTAGGT 501
Db 111 AAANGNAATGGATCGATGATACCTCTTTCAGTCAAAATGTCAGGTTNTGCGNCCCCCAT 52
Qy 502 GAGGCCAATCATCTGCGACCAATGTGCTTCAATAGTCTTCTGGAAACCT 552
Db 51 GAACCCAAATCTCCAGAAGAGNGGTGNCNTCAATAGTTTACTGGAATCCT 1

RESULT 10
US-09-862-802-7
; Sequence 7, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
```

```
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 279...992
; OTHER INFORMATION: protein coding sequence
; NAME/KEY: misc_feature
; LOCATION: 1348
; OTHER INFORMATION: poly A addition motif
US-09-862-802-7

Query Match 21.4%; Score 134; DB 9; Length 1418;
Best Local Similarity 61.4%; Pred. No. 3.3e-31;
Matches 253; Conservative 0; Mismatches 150; Indels 9; Gaps 2;

Qy 222 AGCTGGGGATGTGCCAGCTTCTTGGAAAGTCATTTGGTTCAGTTGCTACTTTCATTTTC 281
Db 584 AGTCTGGAGCTGTGCCCAAGGATTTGGAGGCTATTTTGGTTCCCACTGCTACTTGGTTCC 643
Qy 282 CAGTGA-----AGAGAAGTTTGGTCTAAGAGTGAGCAGAACTGTGTGAGATGGGAGC 335
Db 644 CACAGTTTCTTCATCAGCATCTTTGGAACAAGAGTGAAGAGAACTGCTCCCGCATGGTGC 703
Qy 336 ACATTTGGTTGTCTCAACACAGACGAGCAGAAATTTTCATTTGTCACAGCAGTGAATCA 395
Db 704 TCATCTAGTGGTGATCCAAAGCCAGGAGCAGAGATTTTCATCAGTGGGATCTTGGACAC 763
Qy 396 GTCATTTTCTTATTTTCTGGGGCTTTCAGACCCCAAGGTAATAATAATTGGCAATGGAT 455
Db 764 TCATGCTGCTTATTTTATAGGTTTGGGATAC---AGGCCATCGGCAATGGCAATGGGT 820
Qy 456 TGATAGACACCTTATGAGAAAAATGTCAGATTTTGGGACCTAGGTGAGCCCAATCATTC 515
Db 821 TGATAGACACCACTATGAAGAAAGTATCACATTTGGCCAAATGGTGAAGCCAGCAGTGG 880
Qy 516 TGCAGAGCAATGTGCTTCAATAGTCTTCTGGAACCTACAGGATGGGCTGGAATGATGT 575
Db 881 CAATGAAAATGTGCTACAATAATTTACCGTTGGAAGACTGGATGGGCTGGAACGATAT 940
Qy 576 TATCTGTGAACTAGAGGAATTCATATGTGAGATGAATAAGATTTTACCTA 627
Db 941 CTCTTCAGTCTTTAAACAGAAAGTCAGTTTGTGAGATGAAGAAATAAACTTA 992

RESULT 11
US-09-833-381-1339
; Sequence 1339, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1339
; LENGTH: 758
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-833-381-1339

Query Match 19.2%; Score 120.6; DB 10; Length 758;
```

Best Local Similarity 52.5%; Pred. No. 3.2e-27;  
Matches 321; Conservative 0; Mismatches 274; Indels 16; Gaps 2;  
QY 17 AACCTCAAGTACAGAGAAAAGAGC-TGGTTGTCCCTGAGACTCTGGTCTGTGGCTGGG 75  
Db 18 AACCTCAAGTAAACTGAAGAGGATGCATCCAGCTGTCTCGGTTATTGCTGTA 77  
QY 76 ATTTCCATTGACCTCCTCAGTCTGTCTTCAATTTGTAGCTGTGTAGTAACCTTACCATT 135  
Db 78 GTTTTCATCTTACTCTCAGTGTCTGTTTATTGCAAGTCTTTTGGTGACTCATCAAC 137  
QY 136 ACATATGTGAACCTGCAAAAGGCTGTCTGAACACTACACTATATCAAGTCTCACC 195  
Db 138 TTTTCACGCTGTAGAGAGGCACAGGAGTGCACAAGTTAGAGCACCATGCAAGCTCAA 197  
QY 196 TGTCTCAGT-----AAGGGACAAAGGTGCCAGCTGGGATGTTGCCCA 240  
Db 198 TGCATCAAGAGAAATCAGAACTTAAATACATTAAAGGAGCACCTGGAACGTGTGCT 257  
QY 241 GCTTCTTGGAAAGTCAATTTGGTTCCAGTTGCTACTTCAATTTCCAGTGAAGAGAGTTGG 300  
Db 258 ATTGACTGGAGAGCCTTCCAGTCCAACTGCTATTCTCTTACTGACAAAGACGTGG 317  
QY 301 TCTAAGACTGACGACAACTGTTGAGATGGGACACATTTGGTTGTTCACACAGAA 360  
Db 318 GCTGAGATGAAAGGAAGTGTTCAGGGATGGGGCCCATCTGATGACCATCAGACGGAA 377  
QY 361 CGACAGCAGAAATTTCAATGTCACGACGCTGAATGAGTCATTTTCTTATTTCTGGGCTT 420  
Db 378 GCTGAGCAGAACTTATTATTTCAGTTTCTGATAGACGGCTTCTCTATTTCTTGGACTT 437  
QY 421 TCAGACCCCAAGGTAATAATTTGCAATGGATGATGATGATGATGATGATGATGATGAT 480  
Db 438 AGAGATGAGATGCAAAAGTTCAGTGGGTTGGGTGGACACAGAGCCATTTAAACCCACG 497  
QY 481 GTCAGATTTTGGCACTAGGTGAGCCCAATCATTTCTGCAGCAATGTGCTTCAATAGTC 540  
Db 498 AGAGATTTCTGGCATTAAGAAATGAACCCGACAACTCTCAGGAGAAACATGTTGTTCTT 557  
QY 541 TTCTGGAACCTTACAGGTGGGGTGGAAATGATGTTATCTGAACTAGAAAGAAATTC 600  
Db 558 GTTTATAACCAAGATAAATGGCCCTGGAATGATGTTCTTCTAATTTGAAGCAAGTAGG 617  
QY 601 ATATCTGAGAT 611  
Db 618 ATTTGTAAGT 628

RESULT 12  
US-09-965-529-40  
; Sequence 40, Application US/09965529  
; Publication No. US20020182671A1  
; GENERAL INFORMATION:  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: BURFORD, Neil  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LU, Dyrung Aina M.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0731 USA  
; CURRENT APPLICATION NUMBER: US/09/965,529  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315  
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PERL Program  
; SEQ ID NO 40  
; LENGTH: 968  
; TYPE: DNA

; ORCANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1521513CB1  
US-09-965-529-40  
Query Match 18.7%; Score 117.4; DB 9; Length 968;  
Best Local Similarity 54.4%; Pred. No. 3.5e-26;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;  
QY 17 AACCTCAAGTACAGAGAAAAGAGGCTGTGTGCC---TGAGACTCTGTGGTCTGTGGCTG 73  
Db 184 AACACAATGCACAGAGAGAGGATGCTTCTCTCCCAATGTCTTATGAGACTGTGCTG 243  
QY 74 GGATTTCATTTGCACCTCCCTCAGTGTGCTTTCATTTGTGAGCTGTGTAGTAACCTTACCATT 133  
Db 244 GGATCCCATCTCTATTCTCAGTGTGCTTTTCATCACCAGATGTGTGTGACATTTCCGA 303  
QY 134 TTACATATGTTGNAACCTGGCAAAAGGCTGTCTGAACTACACTCATATCATTTCAAGTCTCA 193  
Db 304 TCTTTCAAAACCTGTGTATGAGAAAAGTT---TCAGCTACCTGAGAATTTTCACAGAGCTCT 360  
QY 194 CCTGCTTTCAGTGAAGGGACAAAAGTCCAGCTGGGGATGTTGCCAGCTTCTTGGAAAT 253  
Db 361 CTTGCTACAAATTTATGATCA---GGTTCACTCAAGAAATTTGTTCATTTGAACCTGGAA 417  
QY 254 CATTGGTTCCAGTGTCTACTTCAATTTCCAGTGAAGAGAAAGTTTGGTCTAAGAGTGGAGC 313  
Db 418 ATTTTCAATCCAGCTCTACTTCTTTCTACTGACACCATTTTCTGGGGGTTTAAAGTTTAA 477  
QY 314 AGAAGTGTGTGAGATGGGAGCACATTTGGTTGTGTTCACACAGACAGAGCAGAAAT 373  
Db 478 AGAAGTGTCTGACCATTTGGGGGCTCACTGGTGTATCAACTCACAGGAGGAGCAGGAAT 537  
QY 374 TCATTTCTCAGAGCTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCACACAG 433  
Db 538 TCCTTTCTTACAGAAACCTTAAATGAGAGAGTTTTATTTGAGCTGTGACAGCAGGTTG 597  
QY 434 GTAATAATATTTGGCAATGGATGATGATGATGATGATGATGATGATGATGATGATGATG 493  
Db 598 TCGAGGCTCAGTGGCAATGGTGGAGCGCACACCTTTGACAAAAGTCTCTGAGCTTCTGGG 657  
QY 494 ACCTAGTGTGAGCCCAATCA-----TTCTGACAGACAATGTCTTCAATAGTCTTCTGGA 547  
Db 658 ATGTAGGGAGGCCCAACAACATAGCTACCTGTGAGGAGCTGTGCCACATGAGAGACTCTT 717  
QY 548 AACCTACAGGATGGGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 607  
Db 718 CAACCCAGGCCAAATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 777  
QY 608 AGATG 612  
Db 778 AAATG 782

RESULT 13  
US-09-905-291A-376  
; Sequence 376, Application US/09905291A  
; Patent No. US20020160374A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/905,291A  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 376  
LENGTH: 997  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-905-291A-376

Query Match 18.7%; Score 117.4; DB 9; Length 997;  
Best Local Similarity 54.4%; Pred. No. 3.6e-26;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

QY 17 AACCTCAAGTACAGAGAAAGGCTGGTGTGCC---TGAGACTCTGGTCTGTGGCTG 73  
DB 128 AAACACAATCACAGAGAGAGAGTCTCTCTCCCAAAATGTTCTTATGGACTGTGGCTG 187  
QY 74 GGATTTCCATGTGACTCCTCAGTGTGCTTCATTGTGAGCTGTAGTAACCTTACCAATT 133  
DB 188 GGATCCCATCTTATTTCTCAGTGCCTGTTTCAACACAGAGTGTGTGACATTTCCGA 247  
QY 134 TTACATATGTGAACCTGCGAAAGGCTGTCTGAACCTACACTCATATCATTTCAAGTCTCA 193  
DB 248 TCITTTCAACCTGTGTGATGAGAAAAGTT---TCAGTACTCTGAGAATTTACAGAGCTCT 304

QY 194 CCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGATGTTGCCAGCTTCTTGAAGT 253  
DB 305 CCTGCTACAATATATGATCA---GGTTCACTCAAGAAATGTTGTCATTAAGTGGGAAT 361  
QY 254 CATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAAAGGTTTGGTCTTAAGAGTGGC 313  
DB 362 ATTTTCAATCCAGCTGCTACTTCTTTTCTACTGACACCAATTTCCCTGGCGGTTAACTTTAA 421  
QY 314 AGAACTGTGTGAGATGGGAGCACATTTGGTGTCTTCAACACAGAAAGCAGAGCAAGATT 373  
DB 422 AGAACTGTGTGAGATGGGAGCACATTTGGTGTCTTCAACACAGAAAGCAGAGCAAGATT 481  
QY 374 TCATTGTCTCAGCAGCTGAATGAGTCAATTTCTTATTTTCTGGGGCTTTTCAGACCCACAAG 433  
DB 482 TCCTTTCTCACAAGAACTTAAATGACAGAGATTTTATTTGGAGCTGTACACACAGGTTG 541  
QY 434 GTAATAATAATTTGCAATGGAATGATAAGACACACCTTATGAGAAAATGTCAGATTTTGGC 493  
DB 542 TCGAGGGTCAGTGGCAATGGGTGGAGCGCACACCTTTGACAAAGTCTCTGAGCTTCTGGG 601  
QY 494 ACCTAGGTGAGCCCAATCA-----TTCTGACAGCAATGCTTCAATAGTCTTCTGGA 547  
DB 602 ATGTAGGGAGGCCCAACAACATAGCTACCTGGAGGAGCTGTGCCACCATGAGAGACTCTT 661  
QY 548 AACCTACAGGATGGGGCTGGGAATGATTTATCTGTGAAACTAGAAAGGAATTCATATCTG 607  
DB 662 CAACCCCAAGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTTATTTCCGATTTG 721  
QY 608 AGATG 612  
DB 722 AAATG 726

RESULT 14  
US-09-902-853-376  
Sequence 376, Application US/09902853  
Publication No. US20020192659A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/902,853  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US/09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-853-376

Query Match 18.7%; Score 117.4; DB 9; Length 997;
Best Local Similarity 54.4%; Pred. No. 3.6e-26;
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

QY 17 AACCTCAAGTACAGAGAAAGCGCTGTTGTCCTCTTCCAAATGTTTATGAGACTGTGGTCTGCTG 73
DB 128 AAACAATACAGACAGAGAGATGCTCTCTTCCAAATGTTTATGAGACTGTGGTCTG 187
QY 74 GSAATTCATTGCACTCTCTGAGTCTGCTTCAATGTTGAGCTGTGTAGTAACCTTACCAAT 133
DB 188 GGATCCCATCTTATTTCTCAGTGCCTGTTTCATCCACAGATGTTGTGACATTCGCA 247
QY 134 TPACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATATCAATCAAGTCTCA 193
DB 248 TCTTTCAAAACCTGTGATGAGAAAAGTT--TCAGCTACCTGAGCAATTCACAGAGCTCT 304
QY 194 CTTGCTTCAGTGAAGGACAAGGTGCGAGCTGGGATGTTGCCAGCTTCTTGGAGT 253
DB 305 CTGCTACAATATGAGTCA--GGTTCAGTCAAGAATGTTGTCTCAATGAACTGGGAAT 361
QY 254 CATTTGGTTCCAGTTGCTTACTTCTTCCAGTGAAGAGAGGTTTGTCTAAGAGTGCAGC 313
DB 362 ATTTTCAATCCAGCTGCTACTCTTTCTACTGACACCAATTTCTCTGGCGTTAAGTTTAA 421
QY 314 AGAATGTTTGGAGTGGGACACATTTGGTTGTTTCAACACAGAAGCAGACGAAT 373
DB 422 AGAAGTCTCAGCCATGGGGCTCACCTGGTGGTTATCACTCAACAGAGGAGCAGGAAT 481
QY 374 TCATTGTCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGCTTCAGACCCACAAG 433
DB 482 TCCTTCTACAAGAAACCTAAAATGAGAGAGTTTTTTATTTGGACTCTCAGACCCAGTTG 541
QY 434 GTAATAAATTTGCAATGGATTGATAGACACACCTTATGAGAAAAATGTGAGATTTTGGC 493
DB 542 TCGAGGGTCAGTGGCAATGGGTGACGCCACACCTTTGACAAAACTCTGAGCTTCTGGG 601

QY 494 ACCTAGTGGAGCCCAATCA-----TTTCGACAGCAATGTGCTTCAATAGTCTTCTCGA 547
DB 602 ATGTAGGGAGCCCAACAACATAGTACCTCGAGGACTGTGCCACCATGAGAGACTCTT 661
QY 548 AACCTACAGATGGGCTCGAATGATGTATCTGTCAAACTAGAAGGAATTCATATATG 607
DB 662 CAAACCCAAAGCAAAATGGAAATGATGAACCTGTTCTCTCAATATTTCGGATTG 721
QY 608 AGATG 612
DB 722 AATG 726

RESULT 15
US-09-907-824-376
; Sequence 376, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564

```
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-824-376

Query Match      18.7%; Score 117.4; DB 9; Length 997;
Best Local Similarity 54.4%; Pred. No. 3.6e-26;
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

Qy 17 AACCTCAAGCTACAGAGAAAGAGGCTGGTTGTCCC---TGAGACTCTCTGTCTGTGGCTG 73
Db 128 AAACACAATCCACAGAGAGGATGCTTCCTTCCCAATGTTCTTATGACTGTGTGCTG 187
Qy 74 GGATTTCCATTGGCAGCTCCTCAGTGTCTGTTTCATTTGTGAGCTGTGTAGTAACCTTACCATT 133
Db 188 GGATCCCAATCCTATTCTCAGTGCCTGTTTCATCACCAGATGTTGTGACATTTTCGCA 247
Qy 134 TTACATATGTTGAACCTGGCAAAAGGCTGCTGAACCTACACTCATATCATTTCAAGTCTCA 193
Db 248 TCTTTCAAACTGTGATGAGAAAAAGTT---TCAGCTTACCCTGAGAATTTTCACAGAGCTCT 304
Qy 194 CCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGATGTTGCCAGCTTCTTTGGAAGT 253
Db 305 CCTGCTACAATTATGGATCA---GGTTCAGTCAAGAAATTTGTTCCATTGAACTGGGAAT 361
Qy 254 CATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAGAGTTTGGTCTAAGAGTGAGC 313
Db 362 ATTTTCAATCCAGCTGCTACTTCTTTCTACTGACACCATTTCTCTGGGCGTTAAGTTTAA 421
Qy 314 AGAAGCTGTTGACATGGAGCAGACATTTGGTGTGTTTCAACAGAGACAGACAGAAAT 373
Db 422 AGAAGCTGCTCAGCCATGGGGCTCACCTGGTGGTTATCAACTCACAGGAGCAGCAAT 481
Qy 374 TCATTGTCACAGCTGAATGAGTCAATTTTCTTATTTCTGGGCTTTCAGACCCACAAG 433
Db 482 TCCTTTCTACAGAAACCTAAATGAGAGAGTTTATTTGAGCTGTCAGACAGGTG 541
Qy 434 GTAATAATAATTGGCAATGGATTGATAGACACCTTATGAGAAAAATGTCAGATTTTGGC 493
Db 542 TCGAGGTCAGTGGCAATGGGTGGCGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGG 601
Qy 494 ACCTAGGTGAGCCCAATCA-----TTCTGAGAGCAATGTGTTCAATAGTCTTCTGGA 547
Db 602 ATGTAGGGGAGCCCAACACATAGACTACCTGGAGAGCTGTGCCACCATGAGAGACTCTT 661
Qy 548 AACCTACAGGATGGGCTGGGAATGATGTTATCTGTGAAGCTAGAAAGAAATTCATATGTG 607
Db 662 CAAACCCCAAGGCAAAATTTGGAATGATGATGTAACCTGTTCTCCTCAAATTTATTTTCGGAATTTGTG 721
Qy 608 AGATG 612
Db 722 AAATG 726
```

Search completed: February 17, 2003, 11:51:36  
Job time : 26.9615 secs





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:47:16 ; Search time 602.362 Seconds  
(without alignments)  
16857.906 Million cell updates/sec

Title: US-09-766-511B-52

Perfect score: 627

Sequence: 1 atgatcgaaagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.4	53.0	504	12	BF603486
2	320	51.0	470	12	BF774158
3	310	49.4	971	13	BI696642
c 4	176	28.1	582	9	AI183967
c 5	172.8	27.6	753	9	AI949910
6	167.2	26.7	715	10	AV716802

c	7	166.8	26.6	1013	9	AA868502
	8	166	26.5	582	10	AV716725
	9	151.4	24.1	514	12	BF152383
	10	151.4	24.1	558	9	AI391056
	11	141.8	22.6	561	10	BE650539
	12	140	22.3	631	11	AK007794
	13	139.4	22.2	1200	11	AK014135
	14	135	21.5	448	9	AA446401
	15	135	21.5	666	10	AK242656
	16	134.6	21.5	812	11	AK020363
	17	134	21.4	568	10	BE200183
	18	130.6	20.8	1007	14	BQ715169
	19	130.2	20.8	403	10	AW822082
	20	119.8	19.1	498	10	BE689245
	21	116.8	18.6	889	12	BG164873
	22	116.4	18.6	1032	10	BE311916
	23	108.6	17.3	608	14	BQ002993
	24	106.4	17.0	667	10	BE688920
	25	105.6	16.8	681	12	BF020383
	26	104.6	16.7	334	10	BE690187
	27	103.4	16.5	831	12	BG065735
	28	99.8	15.9	678	12	BF227688
	29	96.6	15.4	559	10	BE120531
	30	94.8	15.1	576	12	BG079007
	31	94.4	15.1	676	13	BI909298
	32	92.6	14.8	738	14	BQ004812
	33	90.4	14.4	408	10	AV711676
	34	89.4	14.3	544	9	AI400631
	35	88.8	14.2	264	9	AV057557
	36	87.4	13.9	432	10	BB635598
	37	87.2	13.9	228	9	AV252043
	38	86.8	13.8	458	10	BE632986
	39	85.8	13.7	291	9	AA423158
	40	81.6	13.0	631	12	BG080418
	41	80.2	12.8	756	13	BI278559
	42	78	12.4	715	17	AQ050619
	43	77.2	12.3	124	9	AI182435
	44	77.2	12.3	648	12	BF393825
	45	77	12.3	513	9	AA418441

ALIGNMENTS

RESULT 1	BF603486	504 bp	mRNA	linear	EST 25-APR-2001
LOCUS	269014	MARC 3BOV	Bos taurus	cdna 5', mRNA sequence.	
DEFINITION	BF603486				
ACCESSION	BF603486				
VERSION	BF603486.1	GI:11701143			
KEYWORDS	EST.				
SOURCE	cow.				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
	Rovidae; Bovinae; Bos.				
	1 (bases 1 to 504)				
REFERENCE	Smith, F.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,				
AUTHORS	Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett				
	, G.L., Heaton, M.P., Laegreid, W., Rohrer, G.A., Chitko-McKown, C.G.,				
	Perlea, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and				
	Keele, J.W.				

Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)

CONTACT: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred



BI696642	BI696642	771 bp	mRNA	linear	EST 18-SEP-2001
LOCUS	603348532F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5376247 5',				
DEFINITION	mRNA sequence.				
ACCESSION	BI696642				
VERSION	BI696642.1 GI:15659271				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	1 (bases 1 to 771)				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M1957 row: p column: 08 High quality sequence start: 5 High quality sequence stop: 777.				
FEATURES	Location/Qualifiers				
source	1..771				
	/organism="Mus musculus"				
	/strain="FVB/N-3"				
	/db_xref="taxon:10090"				
	/clones="IMAGE:5376247"				
	/clone_lib="NCI_CGAP_Mam2"				
	/tissue_type="tumor_biopsy sample"				
	/dev_stage="5 months"				
	/lab_host="DH10B"				
	/note="organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"				
BASE COUNT	258 a 222 c 229 g 262 t				
ORIGIN					
	Query Match 49.4%; Score 310; DB 13; Length 971;				
	Best Local Similarity 73.8%; Pred. No. 2.7e-85;				
	Matches 450; Conservative 0; Mismatches 150; Indels 10; Gaps 4;				
QY	10 GAGCAGAACCTCAAGTACAGAGAAAGAGCGTCGGTTGCCTCAGACTCTGGTCTG 69				
Db	149 GTGCAGGAAGACAAATCCCAAGGGAAGGAGTCTCTGGACCCTGAGACTCTGGTCA 208				
QY	70 GCTGGGATTTCCATTGCGACTCCTCAGTCTGCTTCAITGTGAGCTGTGTACTAACTTAC 129				
Db	209 GCTGTGATTTCCATGCTTACTCTTGACTACCTCTTTTCATTGCGAGCTGTGTGCTGACTTAC 268				
QY	130 CATTTTACATATGGTGAACACTGGCAAAAGCGTCTGCACTACATCATATCATTTCAACT 189				
Db	269 CAATTTATATTGGACCGACCCAGTAGAGACTATATGAACCTTCACACATACCATTTCCAGT 328				
QY	190 CTCACCTGCTTCAGTGAAGGACAAGGTGCCAG-----CCTGGGGATGTTGCCCCAGCT 243				
Db	329 CTCACCTGCTTCAGTGAAGGACTATGGTGTGAGAAAATCTGGGATGCTGCCCAAT 388				
QY	244 TCTTGAAGTCAATTTGGTTCAGTTGCTACTTCAITTCAGTGAAGAGGTTTGGTCT 303				
Db	389 CACTGGAAGTCAATTTGGTTCAGTCTACCTCAITTTCTACCAAGGAGAACTCTGTGAGC 448				
QY	304 AAGAGTGACACAACCTGTTTCAG-ATGGGACACATTTGGTTGTCTCAACACAGACAGC 362				
Db	449 ACCAGTGAGCAACACTGTGTTTCAGATGGGGCTCATCTGGTGGTGATCAATACTGAAGC 508				
QY	363 AGAGCAGAAATTTCAITGTGCCAGCAGCTGAATGAGTCAITTTCTTAATTTCTGGGCTTTC 422				

```
QY 226 TGGGATGTTGCCAGCTCTCTTGAAGTCAATTTGGTTCCAGTTGCTACTTCAATTTCCAGT 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTCCTACTTATTTCTACT 520
QY 286 GAAGAGAAGGTTTGGCTTAAGACTGAGCAGAACTGTCTGAGATGGGAGCACATTTGGTT 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 GGGATGCAATCTTGGACTAGAGTCAAAAGAACTGTTCTGTGATGGGGGCTGATCGGTG 460
QY 346 GTGTTCAACACAGACAGACAGAGATTTTCATTTGTCAGCAGCTGAATGAGTCAATTTCT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 GTGATCAACACAGGGAAGACAGGATTTTCATTCATCAGAACTCGAAGAAATTTCTTCT 400
QY 406 TATTTTCTGGGGCTTCCAGACCCACACAGGTAATAATTTGGAATGGATGTAAGACA 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 TATTTTCTGGGGCTGTCAGATCCAGGGGTCGCGGACATTTGGCAATGGTGTGACCCAGACA 340
QY 466 CTTTATCAGAAATGTC-----AGATTTGGCACCTAGGTAGGCCCAA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 CCATACAAATGAAATGTCAGCTGAGTATAGAATGAGATTTGCGCCTCAGGTGAACCCAA 280
QY 510 TCATTTGTCAGAGCAATGTCCTTCAATAGTCTTCTGGAACCTACAGGATGGGGCTGGAA 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 TAACCTTGATGAGGCTGTGCGGATATAAATTTCCGTTCTTCAGAGAAATGGGGCTGGAA 220
QY 570 TGATGTTATCTGGAACACTAGAGGAATTCATATATGTCGAGATGAATTAAGATTTACCTA 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 TGACATTCACCTGTCATGTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 162

RESULT 5
AI949910/c
LOCUS      AI949910              753 bp      mRNA      linear      EST 08-MAR-2000
DEFINITION wq04a07.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2470260 3'
            similar to SW:LECI_MOUSE P24721 ASIALOGLYCOPROTEIN RECEPTOR 2 ; ,
            mRNA sequence.
ACCESSION  AI949910
VERSION     AI949910.1  GI:5742220
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 753)
AUTHORS   cDNA Library Prepared by: M. Bento Soares, Ph.D.
            DNA Sequencing by: Greg Lennon, Ph.D.
TITLE     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            DNA Sequencing by: Greg Lennon, Ph.D.
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1065 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 449.
            Location/Qualifiers
                source
                    1..753
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2470260"
                        /clone_lib="NCI_CGAP_Kid12"
                        /tissue_type="p2 pooled tumors (clear cell type)"
                        /lab_host="DH10B"
                        /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
                        Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                        prepared, and ss circles were made in vitro. Following HAP
```

purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "

```
BASE COUNT      230 a      171 c      138 g      212 t      2 others
ORIGIN

Query Match      27.6%; Score 172.8; DB 9; Length 753;
Best Local Similarity 65.4%; Pred. No. 1.6e-42;
Matches 268; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 221 CAGCCTGGGGATGTTGCCAGCTTCTTGGAACTCATTTGGTTCCAGTTGCTACTTTCATTT 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 CAGGCTGGCAGTTGTGCCCAAGAATTCNGGTGTCATTTAGTTCCAACGCACTTTTATTT 678
QY 281 CCAGTGAAGAGAAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTTGAGATGGGACACATT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 CTACTGTAATCAGCATCTGGCAAGACAGTGAAGGACTGTGCTAGAATGNAGGCTCACC 618
QY 341 TCGTTGTGTTCAACACAGACAGCAGAGCAATTTTCATTTCCAGCAGCTGAATCAGTCAT 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 TGCTGGTGATAAACAACCTCAAGAAGCAGAGGATTTTCATCTTCCAGAATCTGCAAGAAGAA 558
QY 401 TTTCTTATTTTCTGGGGCTTTTCAGACCCACAAAGGTAAATAATAATTGGCAATGGATTGATA 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 CTGCTATTTTGTGGGGCTCTCAGATCCAGAAGGTCCAGGACATTTGGCAATGGGTTGATC 498
QY 461 AGACACCTTATGAGAAATAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCATTTCTGCAG 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 AGACACCAACAATAAGAAATTTCCACATTTCCGATCCAGCTGAGCCAGTGATCCCAATG 438
QY 521 AGCAATGCTGTTCAATAGTCTTCTGGAA---ACCTACAGATGGGCTGGAATGATGTTA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 AGCGTGGTGTGCTGCTAAATTTTCGTAATCACCACAAAGATGGGGTGGAAATGATGTTA 378
QY 578 TCTGTGAAACTAGAAGGAATTTCAATATGTGAGATGAATAAGATTTTACCTA 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 ATTGCTTTGGTCTCAAAGGTCAGTTTGTGAGATGATGAAGATCCACTTA 328

RESULT 6
AV716802
LOCUS      AV716802 DCB Homo sapiens cDNA clone DCBAKB04 5', mRNA sequence.
DEFINITION AV716802
ACCESSION  AV716802
VERSION     AV716802.1  GI:10813954
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 715)
AUTHORS   Xu,X., Gu,J., Liu,F., Ou,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,
            G., Cheng,Z. and Han,Z.
TITLE     Homo sapiens cDNA DCB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zequang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
            Location/Qualifiers
                source
                    1..715
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="DCBAKB04"
```



REFERENCE  
1 (bases 1 to 558)  
AUTHORS  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Pe-  
nnington, G., Gibbs, M., Dardik, D., Haines, V., and Schalk,  
D.

E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 The WashU-NCI Mouse EST Project 1999  
 Unpublished (1999)  
 Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 This read is a RESEQUENCE of a previously sequenced mouse clone  
 This read has been verified (found to hit its original self in the correct orientation)  
 MGI:220241  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 491  
 POLYA-NO.

#### FEATURES

source  
 1..558  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:348441"  
 /dev\_stage="19.5 dpc total fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTCGGAGCGCGCATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."  
 171 a 108 c 129 g 150 t

Query Match 24.1%; Score 151.4; DB 9; Length 558;

Best Local Similarity 64.1%; Pred. No. 6.5e-36;

Matches 262; Conservative 0; Mismatches 141; Indels 6; Gaps 2;

QY 222 AGCTGGGATGTGCCAGTCTTGGAGTCATTGGTCCAGTCTGCTACTTCATTC 281  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 76 AGTCTGGACTGTGCCAAAGATGGNAGCCGTTGGTTCCTACTCTACTTCACCTC 135  
 QY 282 CAGTGAAGAAGGT---TTGTCTTAAGAGTGAGCAGAACTGTGTGAGATGGGAGCACA 338  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 136 AACTGACTTGGTGGCATCTTGAATGAGAGTAAGGAGAAGTCTTCCACATGGGTGCTCA 195  
 QY 339 TTGTGTTGTTCAACAGACAGCAGCAGAAATTCATTGTCAGCAGCTGAATGAGTC 398  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 196 TCTGTTGTTGATCCACAGCCAGGAGAACAGGATTCATCACTCGGATCTCGGACATGG 255  
 QY 399 ATTTCTTATTTCTGGGCTTTTCAGACCCACAAAGTAAATAATATGCGCAATGATTC 458  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 256 TACTGCTATTTATAGGACTTTCAATCC---AGTGATCAACAATGCGCAATGATTCGA 312  
 QY 459 TAAGACACCTTATGAGAAAATGTCAGATTTTGGCACCTAGTAGGCCCAATCATTCGTC 518  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 313 TCAGACACCGTACGATGATTAATACCACATCTTGGCACAAGGTAGCGCTAGCAGTGACA 372  
 QY 519 AGACCAATGTGCTCAATAGTCTTCTGGAACCTACAGATGGGGCTGGAATGATGTTAT 578  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 373 TGAACAGTGTATTATAATAAATCATCTGCAGAGTACTGGGCTGGAGTGATATCCC 432  
 QY 579 CTGTGAACACTAGAGGAATTCATATGTGAGATGAATTAAGATTTACCTA 627  
 || || || || || || || || || || || || || || || || || || || || ||  
 Db 433 TTGCAGTGATAACAGAACTCAATTTGGCATGTGAAAAAATAATACCTA 481

#### RESULT 11

BE650539

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BE650539 561 bp mRNA linear EST 06-SEP-2000  
 UI-M-BH2.3-aqf-f-04-0-UI.r1 NIH\_BMAP\_M\_S3.3 Mus musculus cDNA clone  
 UI-M-BH2.3-aqf-f-04-0-UI 5', mRNA sequence.

BE650539 1 GI:9976363

EST.  
 house mouse.  
 Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 561)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene  
 discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: msStem@nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone Distribution:

Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It

should be noted that Bento Soares is generating a small number of

additional specialized non-redundant arrays of BMAP cDNAs whose

availability will be considered under appropriate and limited

collaborative arrangements

Seq primer: M13 Reverse.

Location/Qualifiers

1..561

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-BH2.3-aqf-f-04-0-UI"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; The

NIH\_BMAP\_M\_S3.3 library is a subtracted library of a

series, ultimately derived from a mixture of individually

tagged normalized libraries from ten regions of the mouse

brain (cerebellum, brain stems, olfactory bulbs,

hypothalamus, cortex, amygdala, basal ganglia, pineal

gland, striatum, hippocampus) after a series of

subtractions to reduce the representation of cDNAs from

which ESTs had already been generated. The following

serially subtracted libraries were generated in this

process: NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1.

The subtracted library (NIH\_BMAP\_M\_S3.3) was constructed

as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S2

clones from which 3' ESTs had been derived was used as a

driver in a hybridization with the NIH\_BMAP\_M\_S2 library

single-stranded circles (subtracted library). The remaining

by hydroxyapatite column chromatography, converted to

double-stranded circles and electroporated into DH10B

bacteria (Life Technologies) to generate the

NIH\_BMAP\_M\_S3.3 library. This procedure has been

previously described (Bonaldo, Lennon and Soares, Genome

Research 6:791-806, 1996)"

BASE COUNT 155 a 117 c 144 g 145 t

ORIGIN

Query Match 22.6%; Score 141.8; DB 10; Length 561;

Best Local Similarity 60.9%; Pred. No. 6.5e-33;

Matches 249; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 222 AGCCTGGGAGTGTCCCGAGCTTCTTGCAAGTCATTTGGTTCCAGTGTCTACTTCACTTC 281  
 Db 45 ATCTCGGAGCTGTGTCCAAAGATTTGAAGCCATTTGATTCCTCCACCTGCTACTTCACCTC 104  
 QY 282 CAGTGAAGAGAAGGTTGGTCTAAGAGTGAAGCAGCACTGTGTTCAGATGGGAGCACATTT 341  
 Db 105 CCGTGACACTGCATCTCGAGTGAAGTGAAGAGAAGTGTCTCCCTCAGCGGGTGTCTCACT 164  
 QY 342 GGTGTGTTCAACAGAGAGCAGCAAGATTTTCATGTCAGCAGCAGTGAATGAGTCATT 401  
 Db 165 GCTGTGATCCAGAGCCAGGAAGAGCAGGATTTTCATCCACCAACTCTCGAACCCCTCGTGC 224  
 QY 402 TTCTATTATTTCTGGGCTTTTACAGACCCACAGGTAATAATATGGCAATGGATGATAA 461  
 Db 225 TCGTTATATGTGGGCTCTCAGATCCAAAGGCCATGCAGATGGCAGTGGTGTGATCA 284  
 QY 462 GACACCTTATGAGAAATGTCAGATTTTGGCACCCTAGTGAAGCCCAATCATTTGCGAGA 521  
 Db 285 GACACCATATGATCAAAATGCCACATCTCTGGCACTCAGATGAACCCAGTGGCAACACTGA 344  
 QY 522 GCATGTGCTTCAATAGTCT---TCTGGAACCTACAGATGGGCTGGAATGATGTTAT 578  
 Db 345 ATTTGTGTGTGCTAAGTTATATCAACCGTTAAGAGTGGGCTGGAGTGTGCGCCC 404  
 QY 579 CTGTGAACCTAGAAGGAATTCATATGTGAGATGAATAAGATTTTACCTA 627  
 Db 405 TTGTGATGTGATCAGTGTGATTTGTGAGATGAGCGAGCTCTATGTA 453  
 RESULT 12  
 LOCUS AK007794  
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810046124:similar to DENDRITIC CELL IMMUNORECEPTOR, full insert sequence.  
 ACCESSION AK007794  
 VERSION 1  
 KEYWORDS HT: C57BL/6J  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 REFERENCE 3  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE 4  
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tomita, M., Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombartts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660  
 PUBMED 11217851

5 (bases 1 to 631)  
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to R0 = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGGCGCGCAATTAATTCGCAATTAATTAATTAATTCGCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES  
 Location/Qualifiers  
 1..631  
 /organism="Mus musculus"  
 /db\_xref="FANTOM\_DB:1810046124"  
 /db\_xref="MGI:1901987"  
 /db\_xref="taxon:10090"  
 /clone="1810046124"  
 /sex="male"  
 /tissue\_type="pancreas"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="10 day old"  
 94..624

CDS



/note="data source:SPTR, source key:Q9QZ15, evidence:ISS  
putative  
similar to DENDRITIC CELL IMMUNORECEPTOR"  
/codon\_start=1  
/protein\_id="BAB25260.1"  
/db\_xref="GI:12841568"  
/db\_xref="MGD:MGI:1917060"  
/translation="MYOERQLOQKAVSWLSRLWSAAVISILLSTCFIASCVDKVMWSC  
CPDKWLFGSHCYLVPTVFSSASWNKSENCRMGAHLVVIHQREODFTIGLIDHA  
AFYGLWDTCBQWQVODTPYEESTVFWHNGEPSSDNKCKVTYVYRRNIGWNDIS  
CNLKQKSVCKMKINL"  
BASE COUNT 174 a 129 c 157 g 171 t  
ORIGIN

Query Match 22.3% Score 140; DB 11; Length 631;  
Best Local Similarity 60.6%; Pred. No. 2.5e-32;  
Matches 268; Conservative 0; Mismatches 165; Indels 9; Gaps 2;  
QY 192 CACCTGCTTCAGTGAAGGACAAAGTCCAGGCTGGGGATGTGCCAGCTTCTTGGAA 251  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 183 CACCTGTTTCATTCGAGCTGTGTGGACAAAGTCTGGAGCTGTGCCAAAGCATTTGGAA 242  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 252 GTCATTTGGTCCAGTTCCTACTTCATTTCCAGTGA-----AGACAAGCTTTGGCTTAA 305  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 243 GCTATTTGGTCCACCTGCTACTTGGTCCACAGTTTTTTCATCAGCATCTTTGGAA 302  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 306 GAGTGAGCAACCTGCTTGAGATGGGACGACATTTGGTTGCTTCAACACAGACAGCA 365  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 303 GAGTGAGAGACATGCTCCGCGATGGTGTCTATCTAGTGGTGATCCATGACGAGGAG 362  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 366 GCAGAATTTCATGTCAGGAGTGAATGAGTCAATTTCTTATTTCTGGGGCTTTTTCAGA 425  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 363 GCAGGATTTTCATCCTACTGGGATCTTGGACATTCATGCTGCTTATTTATAGGTTGTGGGA 422  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 426 CCCACAGGTAATAATTTGGCAATGGATTGTAAGACACCTTATGAGAAAATGTCAG 485  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 423 TAC----AGCGCATCGGCAATCGCAATGGTGTGATGATGATGATGATGATGATGATGAT 479  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 486 ATTTTGGCAGCTAGTGCAGCCCAATCATCTTCGAGACCAATGCTTCAATAGTCTTCTG 545  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 480 ATTTGCGCAACATGTCAGCCAGCAGTGCATGCAATGAAATGTTTACAGTATATTACCG 539  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 546 GAAACCTACAGGATGGGGCTGGATGATGTTATCTGTGAACATGAAGGAATTCATATG 605  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 540 TCGGAATATTGGATGGGCTGGAATGATATCTCTTGAATCTTAAACAGAAAGTCAGTTG 599  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 606 TGAGATCAATAAGATTTTACCTA 627  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 600 TCAGATGAAGAAAATAAATTA 621  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13  
AK014135  
LOCUS AK014135 1200 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110037K17:similar to DENDRITIC CELL IMMUNORECEPTOR, full insert sequence.  
ACCESSION AK014135  
VERSION AK014135.1 GI:12851811  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
REFERENCE  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
2

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

COMMENT

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuura,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyama,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,K.I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851  
5 (bases 1 to 1200)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Soabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamana,K.I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to



Search completed: February 17, 2003, 11:47:15  
Job time : 607.362 secs

ACCESSION AW242656  
VERSION AW242656.1 GI:6576410  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 666)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -400P from Gibco  
High quality sequence stop: 422.

FEATURES  
source  
1..666  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2691346"  
/clone\_lib="NCI-CGAP\_Kid11"  
/lab\_host="DH10B"  
/notes="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(clones 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 201 a 147 c 124 g 194 t

Query Match 21.5%; Score 135; DB 10; Length 666;  
Best Local Similarity 65.4%; Pred No. 9.2e-31;  
Matches 214; Conservative 0; Mismatches 110; Indels 3; Gaps 1;  
QY 304 AAGAGTGACGAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTGTTCAACACAGAAAGCA 363  
Db 663 AACAGTGACAAGGACTGTGCAATAATATGATGTCCTCCCTTGGTGATATAACACTCAAGAA 604  
QY 364 GAGCAGAAATTTTCATGCTCCAGCAGCTCAATGAGTCATTTTCTTATTTCTGGGGCTTCA 423  
Db 603 GAGCAGAAATTTTCATGCTCCAGCAGCTCAATGAGTCATTTTCTTATTTCTGGGGCTTCA 544  
QY 424 GACCCACAAAGTAATAATTTGCAATGGATTTGATGAAGACCTTATGAGAAAAATGTC 483  
Db 543 GATCCAGTAGTGCAGCAGCATTTGGTTCATCAGACACCATACATGAAAGTTCC 484  
QY 484 AGATTTTGGCAGCTAGGTGAGCCCAATCATCTGCAGAGCAATGCTTCATAGTCTTC 543  
Db 483 ACATTTCTGGCATCCAGCTGAGCTCAGTTCATGATCCCAATGAGCGCTCGCTTGTGCTAAATTTT 424  
QY 544 TGGAA---ACCTACAGGATGGGCTGGAATGATCTTATCTGTGAACACTAGCAAGGAATTC 600  
Db 423 CGTAATACCCCAAGATGGGCTGGATGATGTTAATGCTCTTGTCTCTCAAGGTCA 364  
QY 601 ATATGTGAGATGAATAAGATTACCTA 627  
Db 363 GTTGTGTGAGATGATGAATCCACTTA 337



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:46:11 : Search time 1081.37 seconds  
(without alignments)  
16874.395 Million cell updates/sec

Title: US-09-766-511B-52

Perfect score: 627

Sequence: 1 atgatgcaagagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pi:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	373.2	59.5	1227	6	AX155225	Sequence
2	373.2	59.5	1227	10	AF240357	Mus muscu
3	373.2	59.5	1254	10	BC023008	Mus muscu
4	273.4	43.6	851	9	AF325459	Homo sapi
5	273.4	43.6	1312	6	AX155223	Sequence
6	273.4	43.6	1313	9	AF293615	Homo sapi
7	271.8	43.3	1125	10	AF240358	Mus muscu
8	269.4	43.0	827	6	AX357481	Sequence
9	265.6	42.4	800	6	AX357483	Sequence
10	243	38.8	1125	10	AF240359	Mus muscu
11	229.8	36.7	739	9	AF325460	Homo sapi
12	201.2	32.1	444	6	AX456975	Sequence
13	200.4	32.0	402	6	AX456962	Sequence
14	192	30.6	1133	9	AF109146	Homo sapi
15	192	30.6	1195	6	AX400064	Sequence
16	192	30.6	1271	9	HSA133532	Homo sapi
17	191.2	30.5	615	9	AF200738	Homo sapi
18	190.4	30.4	1305	9	AF067800	Homo sapi
19	181	28.9	1308	9	AF328684	Homo sapi
C 20	176	28.1	582	6	AX456965	Sequence
C 21	172.8	27.6	753	6	AX456967	Sequence
C 22	166.8	26.6	1013	6	AX456966	Sequence
C 23	151.4	24.1	558	6	AX456968	Sequence
24	151.4	24.1	1192	10	BC034893	Mus muscu
C 25	148	23.6	117296	9	AC092865	Homo sapi
C 26	144.4	23.0	109320	9	AC092746	Homo sapi
27	141.8	22.6	561	6	AX456972	Sequence
28	135.6	21.6	1037	10	AF387099	Mus muscu
29	135	21.5	448	6	AX456973	Sequence
C 30	135	21.5	666	6	AX456971	Sequence
31	134.6	21.5	938	10	BC006623	Mus muscu
C 32	134	21.4	568	6	AX456969	Sequence
C 33	134	21.4	1418	10	MM0133533	Mus muscu
C 34	130.2	20.8	403	6	AX456970	Sequence
C 35	127	20.3	918	10	AF061272	Mus muscu
36	121.8	19.4	1973	9	BC032313	Homo sapi
37	120.2	19.2	753	9	AF411850	Homo sapi
38	117.4	18.7	936	9	BC000715	Homo sapi
39	117.4	18.7	968	6	AX083498	Sequence
40	117.4	18.7	997	6	AX375956	Sequence
41	117.4	18.7	2143	9	AB024718	Homo sapi
42	105.2	16.8	110000	2	AC092450_2	Continuation (3 of
C 43	96.6	15.4	559	6	AX456974	Sequence
C 44	96.6	15.4	5577	2	AC107915	Homo sapi
C 45	96.6	15.4	110000	2	AC092450_2	Continuation (3 of

ALIGNMENTS

RESULT 1  
AX155225  
LOCUS  
DEFINITION Sequence 3 from Patent WO0136487.  
ACCESSION AX155225  
VERSION AX155225.1 GI:14536703  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1227)  
REFERENCE Schmitz,J., Dzionek,A. and Buck,D.W.  
AUTHORS Antigen-binding fragments specific for dendritic cells,  
TITLE compositions and methods of use thereof antigens recognized thereby





JOURNAL expressed on monocyte-derived dendritic cells  
MEDLINE Eur. J. Immunol. 31 (9), 2733-2740 (2001)  
PUBMED 21426806  
11536172  
REFERENCE 2 (bases 1 to 851)  
AUTHORS Arce, I., Hernanz, P. and Fernandez-Ruiz, E.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-2000) Biologia Molecular, Hospital de la  
Princesa, Diego de Leon, 62, Madrid 28006, Spain  
FEATURES  
source Location/Qualifiers  
1..851  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12p12.3-p13.2"  
1..851  
/gene="CLECSF11"  
/gene="CLECSF11"  
/note="C-type lectin receptor; DLEC; alternatively  
spliced"  
/codon\_start=1  
/product="dendritic lectin"  
/protein\_id="AAL37358.1"  
/db\_xref="GI:17225337"  
/translation="MVEPEEPQDRKGLWFKLVMSMAVVSILLISVCFVSSVVP  
NFMYSKTVKRLSKLREYQVHPSLTCVMEGDIKEDWSCPTPTWTSQSCYFTSTGQM  
SWTKSOKNCVWGADLVVINTREODEFLQNLKNSYFLGLSDPGRHHRHOMWDTP  
YENVTWHSEPNLDERCAINFRSEEWGNDHCHVPQKSIKMKKIY"  
BASE COUNT 242 a 169 c 192 g 248 t  
ORIGIN  
Query Match 43.6%; Score 273.4; DB 9; Length 851;  
Best Local Similarity 66.7%; Pred. No. 3.1e-70;  
Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;  
QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGGCTGGTGTCC---TGAGA 57  
DB 27 ATGGTGCCCTGAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGTCCAGTTGAAG 86  
QY 58 CTCGTGCTGTGGTGGGATTTCCATTCGACTCCTCAGTGTGCTTTCATTTGAGCTGT 117  
DB 87 GTCTGTGTCATCGCAGTGTATCCATTCGCTCCCTCAGTGTGCTGTTTCACTGTGAGTTCT 146  
QY 118 GTAGTAACCTTACATTTTACATATGCTGAAACTGGCAAAAGGCTGTCTGAACTA----- 171  
DB 147 GTGGTGCCCTCACAATTTATGTATAGCAAACTGTCAAGAGCTGTCAAGTTACGAGAG 206  
QY 172 ---CACTCATATCATTCAGTCTACCTGCTTCACTGAGGAGCAAAAGTGCACGCTGG 228  
DB 207 TATCAACAGTATCATCCAAGCTGACCTGGCTCATGGAAGAAAGGACATAGAAGATTGG 266  
QY 229 GGATGTGGCCAGCTCTTGGGAAGTCAATTTGGTCCAGTGTGCTACITTCATTTCCAGTGA 288  
DB 267 AGCTGTGTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTGTGCTACTTATTTCTACTGG 326  
QY 289 GAGAAGTTTGTCTAAGAGTTCAGCAGAACTGTGTGTAGATGGGAGCACATTTGGTTGTG 348  
DB 327 ATGCAATCTTGGACTAAGAGTCAAAAGAACCTGTCTGTATGGGGCTGTATCTGGTGTG 386  
QY 349 TTCAACACAGAGCAGCAGCAATTTCAATTTGCCAGAGCTGAATGAGTCAATTTCTTAT 408  
DB 387 ATCAACACCCAGGGAAGAACAGGATTTTCATTCATTCAGAACTCTGAAAGAAATTTCTTAT 446  
QY 409 TTTCTGGGCTTTACAGCCCAACAGGTAATAATTTGGCAATGATTAAGACACCT 468  
DB 447 TTTCTGGGCTGTACAGATCCAGGGGTGGCGACATTTGCAATGGTTGACACAGACCA 506  
QY 469 TATGAGAAAATGTACAGTTTGGCACTTAGTGGAGCCCAATCACTCTCGAGAGCAATGT 528  
DB 507 TACAATGAANAATGTACATTTCTGGCACTCAGGTGAACCCAAATACCTTGTATGAGGTTGT 566  
QY 529 GCTTCAATAGTCTTCTGGAAACCTACAGGATGGGCTGGAATGATGTATCTGTGAAACT 588

DB 567 GCATATAAATTTCCGTTCTTTCAGAGAATGGGCTGGAATGACATTCACTGTCTATGTA 626  
QY 589 AGAAGGAATTCATATGTGAGATGAATAAGATTACCTA 627  
DB 627 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 665  
RESULT 5  
AX155223  
LOCUS AX155223 1312 bp mRNA linear PAT 22-JUN-2001  
DEFINITION Sequence 1 from Patent WO0136487.  
ACCESSION AX155223  
VERSION AX155223.1 GI:14536702  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1312)  
AUTHORS Schmitz, J., Dzionek, A. and Buck, D. W.  
TITLE Antigen-binding fragments specific for dendritic cells,  
compositions and methods of use thereof antigens recognized thereby  
and cells obtained thereby  
JOURNAL Patent: WO 0136487-A 1 25-MAY-2001;  
Miltényi Biotech GmbH (DE)  
FEATURES  
source Location/Qualifiers  
1..1312  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="BDCA-2 cDNA sequence"  
BASE COUNT 374 a 301 c 297 g 340 t  
ORIGIN  
Query Match 43.6%; Score 273.4; DB 6; Length 1312;  
Best Local Similarity 66.7%; Pred. No. 3.2e-70;  
Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;  
QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGGCTGGTGTCC---TGAGA 57  
DB 491 ATGGTGCCCTGAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGTCCAGTTGAAG 550  
QY 58 CTCGTGCTGTGGTGGGATTTCCATTCGACTCCTCAGTGTGCTTTCATTTGAGCTGT 117  
DB 551 GTCTGTGTCATCGCAGTGTATCCATTCGCTCCTCAGTGTGCTTTCATTTGAGTTCT 610  
QY 118 GTAGTAACCTTACATTTTACATATGCTGAAACTGGCAAAAGGCTGTCTGAACTA----- 171  
DB 611 GTGGTGCCCTCACAATTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAGTTACGAGAG 670  
QY 172 ---CACTCATATCATTCAGTCTACCTGCTTCACTGAGGAGCAAAAGTGCACGCTGG 228  
DB 671 TATCAACAGTATCATCCAAGCTGACCTGGCTCATGGAAGAAAGGACATAGAAGATTGG 730  
QY 229 GGATGTGGCCAGCTCTTGGGAAGTCAATTTGGTTCAGTGTGCTACITTCATTTCCAGTGA 288  
DB 731 AGCTGTGTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTGTGCTACTTATTTCTACTGG 790  
QY 289 GAGAAGTTTGTCTAAGAGTTCAGCAGAACTGTGTGTAGATGGGAGCACATTTGGTTGTG 348  
DB 791 ATGCAATCTTGGACTAAGAGTCAAAAGAACCTGTCTGTATGGGGCTGTATCTGGTGTG 850  
QY 349 TTCAACACAGAGCAGCAGCAATTTCAATTTGCCAGAGCTGAATGAGTCAATTTCTTAT 408  
DB 851 ATCAACACCCAGGGAAGAACAGGATTTTCATTCATTCAGAACTCTGAAAGAAATTTCTTAT 910  
QY 409 TTTCTGGGCTTTACAGCCCAACAGGTAATAATTTGGCAATGATTAAGACACCT 468  
DB 911 TTTCTGGGCTGTACAGATCCAGGGGTGGCGACATTTGCAATGGTTGACACAGACCA 970  
QY 469 TATGAGAAAATGTACAGTTTGGCACTTAGTGGAGCCCAATCACTCTCGAGAGCAATGT 528  
DB 971 TACAATGAANAATGTACATTTCTGGCACTCAGGTGAACCCAAATACCTTGTATGAGGTTGT 1030



QY	529	GCTTCAATAGTCTTCTGGAAACCTACAGATGGGCTGGAAATGATGTTATCTGTGAAACT	588
Db	1031	CGGATAATAATTTCCGTTCTTACAGAGATGGGCTGGAAATGACATCTCACTGTCATGTA	1090
QY	589	AGAAGGAATTCATATGTGAGATGAATAAGATTACCTA	627
Db	1091	CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA	1129
RESULT 6			
LOCUS	AF293615	1313 bp	mRNA linear PRI 14-FEB-2002
DEFINITION	Homo sapiens blood dendritic cell antigen 2 protein (BDCA2) mRNA, complete cds.		
ACCESSION	AF293615		
VERSION	AF293615.1	GI:17224597	
KEYWORDS	Homo sapiens.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1313)		
AUTHORS	Dzionek,A., Sohma,Y., Nagafune,J., Cella,M., Colonna,M., Facchetti,F., Gunther,G., Johnston,I., Lanzavecchia,A., Nagasaka,T., Okada,T., Vermi,W., Winkels,G., Yamamoto,T., Zysk,M., Yamaguchi,Y. and Schmitz,J.		
TITLE	BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction		
JOURNAL	J. Exp. Med. 194 (12), 1823-1834 (2001)		
MEDLINE	21615135		
PUBMED	11748283		
REFERENCE	2 (bases 1 to 1313)		
AUTHORS	Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J., Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.		
TITLE	Cloning and characterization of blood dendritic cell antigen 2 (BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 1313)		
AUTHORS	Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J., Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-AUG-2000) Miltenyi Biotec GmbH, Friedrich-Ebert-Str. 68, Bergisch Gladbach 51429, Germany		
FEATURES	Location/Qualifiers		
source	1..1313		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/cell_type="plasmacytoid dendritic cell; DC-2"		
	1..1313		
	/gene="BDCA2"		
	492..1133		
	/gene="BDCA2"		
	/notes="C-type lectin"		
	/codon_start=1		
	/product="blood dendritic cell antigen 2 protein"		
	/protein_id="AAL37036.1"		
	/db_xref="GI:17224598"		
BASE COUNT	374 a	302 c	297 g
ORIGIN	340 t		
Query Match	43.6%	Score 273.4;	DB 9; Length 1313;
Best Local Similarity	66.7%;	Pred. No. 3.2e-70;	
Matches 426;	Conservative	0; Mismatches 201;	Indels 12; Gaps 2;
QY	1	ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGCTGGTGTGCC---	TGAGA 57
Db	492	ATGTGCTGAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGTCCAGTTGAAG	551

QY	58	CTCTGGCTGTGGCTGGGATTTCCATTGTCACCTCCCTCAGTCGCTTGGTTCATTGTCGAGCTGT	117
Db	552	GTCTGGTCCAGCGCAGTCGTATCCCATCTTGGCTCCCTCAGTCGCTGCTTCTCACTGTGAGTTCT	611
QY	118	GTAGTAACCTTACCATTTTACATATGTTGAACTGGCAAAAGGCTGTCTCAACTA-----	171
Db	612	GTGGTGCCTCAACAATTTTATGTATACAAAACCTGTCAAGAGGCTGTCCAAGTTACGAGAG	671
QY	172	----CACTCATATCATTCACAGTCTCACCTGCTTCAGTGAAGGAGCAAAAGTCCAGCGCTGG	228
Db	672	TATCAACAGTATCATCAAGCCTGACCTGCTCATGGAAGAGAAAGACATAGAAGATTGG	731
QY	229	GGATGTGGCCAGCTTCTTGGGAAGTCATTGTTCCAGTTGCTACTTTCATTCCACTGAA	288
Db	732	AGCTGCTGCCCAACCCCTTGGACTTCATTTTCAGTCTAGTTGCTACTTTTATTTCTACTGGG	791
QY	289	GAGAAGTTTGGTCTAAGAGTGAAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTG	348
Db	792	ATGCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGGTG	851
QY	349	TTCACACAGAGCAGCAGAGAGAAATTTTCATTGTCACGACGCTGAATGAGTCATTTCTTAT	408
Db	852	ATCAACACCAGGGAGACAGAGGATTTTCATCATTCAGAATCTGAAAAAGAAATTTCTTCTAT	911
QY	409	TTTCTGGGCTTTCAGACCCACAAAGTAATAATAATTGGCAATGGATGTGATGAAGACCT	468
Db	912	TTTCTGGGCTGTGAGATCCAGGGGTGGGACATTTGGCAATGGGTGACACACACCA	971
QY	469	TATGAAAAAATGTCCAGATTTTGGCACCTGAGTGGAGCCCAATCATTTCTCGACAGCAATGT	528
Db	972	TACAATGAAATGTCCACATTTCTGGCACTCAGGTGAACCAATTAACCTTGATGAGCGTTGT	1031
QY	529	GCTTCAATAGTCTTCTGGAAACCTACAGATGGGCTGGAATGATGTTATCTGTGAAACT	588
Db	1032	GGCATAAATAATTTCCGTTCTTTCAGAAAGAAATGGGCTGGAATGACATTCACCTGTCTATGTA	1091
QY	589	AGAAGGAATTCATATGTGAGATGAATAAGATTTCACCTA	627
Db	1092	CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA	1130
RESULT 7			
LOCUS	AF240358	1125 bp	mRNA linear ROD 02-MAY-2000
DEFINITION	Mus musculus strain BALB/c dectin-2 beta isoform mRNA, complete cds, alternatively spliced.		
ACCESSION	AF240358		
VERSION	AF240358.1	GI:7677471	
KEYWORDS	Mus musculus.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1125)		
AUTHORS	Artizumi,K., Shen,G.-L., Shikano,S., Ritter,R. III, Zukas,P., Edelbaum,D., Morita,A. and Takashima,A.		
TITLE	Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms		
JOURNAL	J. Biol. Chem. (2000) In press		
REFERENCE	2 (bases 1 to 1125)		
AUTHORS	Artizumi,K. and Takashima,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA		
FEATURES	Location/Qualifiers		
source	1..1125		
	/organism="Mus musculus"		
	/strain="BALB/c"		
	/db_xref="taxon:10090"		
	/cell_line="XS52"		
	/cell_type="dendritic cell"		
	146..673		
CDS			

/note="C-type lectin; type II transmembrane protein; dendritic cell-associated expression; C-type lectin motif; alternatively spliced form of dectin-2 alpha mRNA"  
/codon\_start=1  
/product="dectin-2 beta isoform"  
/protein\_id="AAP67178.1"  
/db\_xref="GI:7677472"

/translation="MVDEROSQGVGVCTVRLWSAAVISMILLSTCFIASCVBKMMGC  
CPNHWKFGSSCYLITSEKFWSTSEQNCVOMGAHLVINTAEQNFITQOLNLSLSY  
FLGLSDPGNKGWQWIDDTFQSNVRWHPNLPPEERCVSVIYWNPSKWGNDFVC  
DSKHSNLCENKKIYL"

BASE COUNT 305 a 246 c 257 g 317 t  
ORIGIN

Query Match 43.3%; Score 271.8; DB 10; Length 1125;

Best Local Similarity 76.6%; Pred. No. 9.4e-70;

Matches 333; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 193 ACCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGCCAGCTTCCTGGGAAG 252  
DB 236 ACCTGCTTCATTCGAGCTGTGTGAAAAAATGTGGGATGTGCCAAATCACTGGGAAG 295  
QY 253 TCATTTGGTTCAGTTCCTACTTCCAGTGAAGAGAGGTTTGCTCTAAGAGTCAG 312  
DB 296 TCATTTGGCTCCAGCTGTACCTCATTTCTACCAAGGAGAACTTCGGAGCACCAGTCAG 355  
QY 313 CAGAACTGTGTGAGATGGGACACATTTGGTTGTTCACACAGAGACAGACAGAAAT 372  
DB 356 CAGAACTGTTCAGTGGGGCTCATCTGGTGTGATCAATCTAAGAGCGGACAGAAAT 415  
QY 373 TTCATTGCTCCAGCTGAATGAGTCATTTCTTATTTCTGGGGCTTTTCAGACCCACAA 432  
DB 416 TTCATCACCACGAGCTGAATGAGTCACCTTCTTACTTCTCGGTGGTCTTCGGATCCACAA 475  
QY 433 GGTAAATAAATGGCAATGGATTCATAGACACCTTATGAGAAAAATGTCAGATTTGG 492  
DB 476 GGTAAATGGCAATGGCAATGGATTCATAGACACCTTCTTCAAGTCAAAATGTGAGGTCCTGG 535  
QY 493 CACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTCTCAATAGTCTCTGGAAAGCT 552  
DB 536 CACCCCATGAACCAATCTCCAGAGAGCGGTGTGTTCAATAGTTTACTGGAATCCT 595  
QY 553 ACAGGATGGGCTGGAATGATGATTTATCTGTGAACCTAGAAGCAATCAATATGTGAGATG 612  
DB 596 TCGAAATGGGCTGGAATGATGATTTCTGTGTAGTAAACAAATCAATATATGTGAAATG 655  
QY 613 AATAAGATTTACCTA 627  
DB 656 AAGAAGATTTACCTA 670

RESULT 8  
AX357481  
LOCUS AX357481 827 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 19 from Patent WO01/2773.  
ACCESSION AX357481  
VERSION AX357481.1 GI:18674538  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.  
TITLE Organic compounds  
JOURNAL Patent: WO 01/2773-A 19 04-OCT-2001;  
Novartis AG (CH)

FEATURES  
source Location/Qualifiers  
1..827  
/organism="Homo sapiens"  
/db\_xref="Taxon:9606"

BASE COUNT 242 a 164 c 188 g 233 t  
ORIGIN

Query Match 43.0%; Score 269.4; DB 6; Length 827;  
Best Local Similarity 66.8%; Pred. No. 4.8e-69;  
Matches 419; Conservative 0; Mismatches 196; Indels 12; Gaps 2;

QY 13 CAGCAACCTCAAGTACAGAGAAAGAGCTGGTCTGCC---TGAGACTCTGGTCTGG 69  
DB 28 CAGAGGCTTCAAGACCCAGAGAAAGACTCTGGTGGTTCAGTTGAAGGTCTGGTCCATG 87  
QY 70 GCTGGGATTTCCATTGACATCCCTCAGTGTGCTTCCATTTGAGTGTGTAGTAACATTAC 129  
DB 88 GCAGTCCGTATCCATCTTGCCTCAGTGTGTTTCTCAGTGTGAGTCTGTGGTGCCTCAC 147  
QY 130 CATTTACATATGTTGAACTGGCAAAAGCTGTCTGAACTA-----CACTCATAT 180  
DB 148 AATTTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGAGAGATCAACAGTAT 207  
QY 181 CATTCAGTCTCACCCTCAGTGAAGGACAAAGTGCACCCCTGGGATGTTGCCCA 240  
DB 208 CATCCAAGCCTGACCTCGCTCATGGAAGGAAAGGACATAGAAGATTGGAGCTGCTGCCCA 267  
QY 241 GCTTCTTGGAAATCATTTGGTTCAGTGTCTACTTTCATTTCCAGTGAAGAGAGGTTTGG 300  
DB 268 ACCCTTGGACTTCATTTCACTGTAGTGTCTACTTTATTTCTACTGGGATGCAATCTGG 327  
QY 301 TCTAAGAGTGAAGCACTGTGTGAGATGGGACACATTTGGTGTGTTCACACAGAA 360  
DB 328 ACTAAGAGTCAAAAGAACTGTCTGTGATGGGGCTGATCTGGTGTGATCAACACCAGG 387  
QY 361 GCAGAGCAGAAATTCATTTGCTCCAGTGTCTACTTTCATTTCCAGTGAAGAGAGGTTTGG 420  
DB 388 GAAGAACAGGATTCATTCATTCAGAAATCTGAAAGAAATTCCTTTATTTCTGGGGCTG 447  
QY 421 TCAGACCCACAAAGTAAATAATTTGGCAATGGATTGATAAGACACCTTATGAGAAAAAT 480  
DB 448 TCAGATCCAGGGGTGGCCACATTTGGCAATGGTTGACGACACACCATACATGAAAT 507  
QY 481 GTCAGATTTTGGACCTTAGGTGAGCCCAATCATTTCTGCGAGCAATGTGTTCAATAGTC 540  
DB 508 GTCACATTTCTGGCCTCAGTGAACCAATTAACCTTGTGAGCGGTGTGCGGATAATAAT 557  
QY 541 TTCTGGAAACCTACAGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGGAATTC 600  
DB 568 TTCGCTTTTCAGAGAAATGGGCTGGAATGACATGACATGCTATGTCACCTCAAGAGTCA 627  
QY 601 ATATGTGAGATGAATAAGATTTACCTA 627  
DB 628 ATTTGCAAGATGAAGAGATCTACATA 654

RESULT 9  
AX357483  
LOCUS AX357483 800 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 21 from Patent WO01/2773.  
ACCESSION AX357483  
VERSION AX357483.1 GI:18674539  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.  
TITLE Organic compounds  
JOURNAL Patent: WO 01/2773-A 21 04-OCT-2001;  
Novartis AG (CH)

FEATURES  
source Location/Qualifiers  
1..800  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 231 a 160 c 183 g 226 t  
ORIGIN



MEDLINE 21426806  
PUBMED 11536172  
REFERENCE 2 (bases 1 to 739)  
AUTHORS Arce,I., Hernanz,P. and Fernandez-Ruiz,E.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-2000) Biologia Molecular, Hospital de la Princesa, Diego de Leon, 62, Madrid 28006, Spain  
FEATURES  
source  
1..739  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
1..739  
/map="12p12.3-p13.2"  
gene  
8..556  
/gene="CLECSF1"  
CDS  
/gene="CLECSF1"  
/note="C-type lectin receptor; DLEC-b; alternatively spliced"  
/codon\_start=1  
/product="dendritic lectin b isoform"  
/protein\_id="AAL37359.1"  
/db\_xref="GI:17225339"  
/translation="MVPEEPQDRVPHNPMYSKVKLSKLREYQYHPSLTCVMEGK  
DIEDWSCPPTWTSFQSCYFISTGMSWTKSKNSVMGADLVVINTREEQDFIION  
LKRNSYFLGLSDPGRRHWQWDQTPYENVTFWHSGEPNNLDERCALINFRSSEW  
GWNDIHCHVPQKSTCKMKKIY"  
BASE COUNT 222 a 145 c 158 g 214 t  
ORIGIN  
  
Query Match 36.7%; Score 229.8; DB 9; Length 739;  
Best Local Similarity 66.9%; Pred. No. 3e-57;  
Matches 346; Conservative 0; Mismatches 162; Indels 9; Gaps 1;  
  
QY 120 AGTAACCTTACCAATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTA----- 171  
Db 37 AGTGCTCACAATTTATGTATAGCAAAACCTGTCAAGAGGCTGTCCAAAGTTACAGAGTA 96  
  
QY 172 -CACATCATATTCACAGTCTCACCTGCTCAGTGAAGGGCAAAAGGTGCCAGCTGGGG 230  
Db 97 TCACAGATATCATCAAGCTGACCTGCGTGTATGGAAGGAAAGGACATAGAAGATTGGAG 156  
  
QY 231 ATGTTGCCAGCTTCTTGAAGTCACTTGGTTTCCAGTTCAGTTGCTACTTTCATTTCCAGTGAAGA 290  
Db 157 CTGCTGCCCAACCCCTTGGACTTCATTTCAGTCTAGTTGTCTACTTTATTTCTACTGGGAT 216  
  
QY 291 GAAGGTTTGGTCTAAGAGTGAGCAACACTGTGTTGAGATGGGAGCACATTTGGTGTGTT 350  
Db 217 GCAATCTGGACTAGAGTCAAAAGCACTGTCTGTATGGGGCTGATCTGGTGGTAT 276  
  
QY 351 CAACACAGAGCAGCAGCAATTTTCATTTGTCAGAGCTGAATGAGTCAITTTCTTAATT 410  
Db 277 CAACACAGGGAAGAACAGGATTTTCATCATTCAGAAATCTGAAGAAATTTCTTCTATT 336  
  
QY 411 TCTGGGGCTTTCAGACCCACACAGGTAAATAATTTGGCAATGGATTGATAGACACCTTA 470  
Db 337 TCTGGGGCTGTCCAGATCCAGGGGTCGCGACATTTGGCAATTTGGTTGACCAGACACCATA 396  
  
QY 471 TGAGAAAATATGCAGATTTGGCACTAGGTGAGCCCAATCATTTCTCAGAGCAATGTGC 530  
Db 397 CAATGAAATATGCATATCTGGCACTCAGGTGACCCCAATACCTTGTATGAGCGTTGTGC 456  
  
QY 531 TTCAATAGTCTTCTGGAACCTTACAGGATGGGGCTGGAATGATGTATCTGTGAAACTAG 590  
Db 457 GATAATAATTTCCGTTCTTTCAGAGAAGATGGGGCTGGAATGACATTCATCTCATGTACC 516  
  
QY 591 AGGGAATTCATATGTGAGATGAATAAGATTTACCTA 627  
Db 517 TCAGAAGTCAATTTTCAAGATGAAGAAGATCTACATA 553  
  
RESULT 12  
AX456975

LOCUS AX456975 444 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 14 from Patent WO0232958.  
ACCESSION AX456975  
VERSION AX456975.1 GI:21715770  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Smolyar,A.  
TITLE Polynucleotide and polypeptide sequences of human dendritic cell immunoreceptors  
JOURNAL Patent: WO 0232958-A 14 25-APR-2002;  
Bayer Aktiengesellschaft (DE)  
FEATURES  
source  
1..444  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 123 a 86 c 101 g 134 t  
ORIGIN  
  
Query Match 32.1%; Score 201.2; DB 6; Length 444;  
Best Local Similarity 68.5%; Pred. No. 9.4e-49;  
Matches 278; Conservative 0; Mismatches 128; Indels 0; Gaps 0;  
  
QY 222 AGCTGGGATGTTGCCAGCTTCTTGGAAATCATTTGGTTCCAGTTGCTACTTCAATTC 281  
Db 39 AGATTGGAGCTGCTGCCAACCCCTTGGACTTCATTTCACTAGTTAGTTGCTACTTTATTC 98  
  
QY 282 CAGTGAAGAGAGGTTTGGTCTAAGAGTGAAGAGTCTGTGAGATGGGAGCACATTT 341  
Db 99 TACTGGATGCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCT 158  
  
QY 342 GGTGTGTTTCAACACAGAAGCAGACAGCAATTTTCATTTGCCAGAGCTGAATGAGTCAAT 401  
Db 159 GGTGTGTTCAACACAGCAGGAGAACAGGATTTTCATTCATTCAGAAATCTCAAAAGAAATTC 218  
  
QY 402 TTCTTATTTCTGGGCTTTTCAGACCCCAAGTAAATAAATTTGGCAATGAGTGAATAA 461  
Db 219 TTCTTATTTCTGGGCTGTTCAGATCCAGGGGTCGCGGACATTTGGCAATGGGTTGACCA 278  
  
QY 462 GACACCTTATGAGAAAATCTCAGATTTGGCACCTAGGTGAGCCCAATCAITTCGCACA 521  
Db 279 GACACCATACAATGAAAATGTCACATTCGGCACTCAGGTGAACCCCAATACCTTGAATGA 338  
  
QY 522 GCAATGTGTTTCAATAGTCTTCTGGAACCTACAGGATGGGCTGGAATGATGTTATCTG 581  
Db 339 GCCTGTGGCGATAATAAATTCGGTTCTTCAGAAAGAAATGGGCTGGAATGACATTCACATG 398  
  
QY 582 TGAACCTAGAGGAATTCATATGTGAGATGAATAAGATTTACCTA 627  
Db 399 TCATGTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 444  
  
RESULT 13  
LOCUS AX456962 402 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 1 from Patent WO0232958.  
ACCESSION AX456962  
VERSION AX456962.1 GI:21715759  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Smolyar,A.  
TITLE Polynucleotide and polypeptide sequences of human dendritic cell immunoreceptors  
JOURNAL Patent: WO 0232958-A 1 25-APR-2002;  
Bayer Aktiengesellschaft (DE)  
FEATURES  
Location/Qualifiers

```
source 1. .402
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 114 a 80 c 96 g 112 t
ORIGIN
Query Match 32.0%; Score 200.4; DB 6; Length 402;
Best Local Similarity 68.7%; Pred. No. 1.6e-48;
Matches 276; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 226 TGGGATGTTGCCAGCTCTTGGAAGTCATTTGGTTCAGTTCCTACTCTCAATTCACGT 285
DB 1 TGGAGCTGCTGCCAACCCCTGGACTTCATTTTCAGTCTAGTTCCTACTTTATTTCTACT 60

QY 286 GAAGAGAAGGTTTGCTCTAAGAGTGAGCAGAACTGTGTGGAGATGGGAGCACATTTGGTT 345
DB 61 GGGATGCAATCTTGGACTTAAGAGTCAAAGAAGCTGTCTGTGATGGGGCTGATCTGGTG 120

QY 346 GTGTTTCACAGACAGACAGCAAGATTTTCATTTGCCAGCAGCTGAATGAGTCATTTTCT 405
DB 121 GTGATCAACACAGGGAAGACAGGATTTTCATTCATTCAGAACTCTGAAAAGAAATTCITCT 180

QY 406 TATTTTCTGGGGCTTTCAGACCCACAGAGTAAATTAATTCGGAATGGATGATGAAGACA 465
DB 181 TATTTTCTGGGGCTGTCTAGATCCAGGGGTCGGCGACATTTGGCAATGGGTTTGACCGACA 240

QY 466 CCTATGAGAAAAATGTCAGATTTTGGCACCTAGGTAGTGAGCCCAATCTCTGCAGAGCAA 525
DB 241 CATACATGAAATGTGCATCTCGGCATTCAGGTGAGTGAACCAATACCTTGATGAGCGT 300

QY 526 TGTGCTTCAATAGTCTTCTGGAACCTACAGGATGGGGCTGGAATGATGTTATCTGCAA 585
DB 301 TGTGCGAATAAATTTCCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACCTGTCAT 360

QY 586 ACTAGAAGAAATCAATATGTCAGATGAATGAATTAAGATTACCTA 627
DB 361 GTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 402

RESULT 14
AF109146 1133 bp mRNA linear PRI 24-OCT-2000
LOCUS Homo sapiens C-type lectin superfamily 6 (CLECSF6), mRNA, complete cds.
DEFINITION
ACCESSION AF109146
VERSION AF109146.1 GI:6502534
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1133)
Richard, M. and Beaulieu, A.D.
A novel C-type lectin expressed in GM-CSF stimulated neutrophils
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 1133)
Richard, M. and Beaulieu, A.D.
Direct Submission
AUTHORS
TITLE Submitted (24-NOV-1998) Dept. of Medicine, Centre Hospitalier de
JOURNAL l'Universite Laval, 2705 Boul. Laurier, Sainte-Foy, Quebec G1V 4G2,
Canada
FEATURES
source
1. .1133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="polymorphonuclear cells"
gene 1. .1133
/gene="CLECSF6"
CDS 95. .808
/gene="CLECSF6"
/note="C-type lectin DB27"
/codon_start=1
/product="C-type lectin superfamily 6"
```

```
/protein_id="AAF14348.1"
/db_xref="GI:6502535"
/translation="MTSEITYAEVRFKNEFKSSGINTASSASKERTAPLKSNTGFPK
LLCASILLFFLLLAISFPIAFVIFQKYSOLLEKTKTELVHTTLECYKKNPVRETA
WSCCPKNKWSFSSNCYFFISTESASWQSEKDCARMEAHLLVINTQEEODFIQNLOEE
SAIFYGLSDPEGRHQWQVDQTPYNESSTFWHPRESDPNERCVVLNFRKSPKRWGWN
DYNCLGPQRSVCVMKHL"
BASE COUNT 332 a 214 c 239 g 348 t
ORIGIN
Query Match 30.6%; Score 192; DB 9; Length 1133;
Best Local Similarity 68.0%; Pred. No. 5.5e-46;
Matches 283; Conservative 0; Mismatches 130; Indels 3; Gaps 1;

QY 215 AGGTGCCAGCTGGGATGTTGCCAGCTTCTTTGGAAGTCATTTGCTCCAGTTGCTACT 274
DB 390 AAGAGACAGCTGGAGCTGTGGCCAAAGAAATTTGGAAGTCATTTAGTTCCAACTGCTACT 449

QY 275 TCATTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGTGAGCAGAACTGTGTGAGATGGGAG 334
DB 450 TTATTCTACTGAATCAGCATCTTGGCAGACAGTCAGAAGGACTGTGCTAGAATGGAGG 509

QY 335 CACATTTGGTTGTTCAACACAGACAGCAGACAGAAATTTTCATTTGCCAGCAGCTGAATG 394
DB 510 CTCACCTGCTGTCATAAACACTCAAGAAGACGAGGATTTTCATCTTCCAGAATCTCAAG 569

QY 395 AGTCATTTCTTATTTCTGGGGCTTTCAGACCCACAAAGCTAATAATTTGGCAATGGA 454
DB 570 AGAAATCTGCTTATTTTGTGGGGCTCTCAGATCCAGAAGTCCAGCAGCATTTGGCAATGGG 629

QY 455 TTGATAAGACACCTTATGAGAAAATGTCAGATTTTGGCACCTTAGTGAGGCCAACATCAT 514
DB 630 TTGATCAGACACCATACAATGAAAGTTCCACATTTCTGGCATCCACGTGAGCCAGTGATC 689

QY 515 CTGCAGAGCAATGTGCTTCAATAGTCTTCTGGAA---ACCTCAGAGATGGGGCTGGAATG 571
DB 690 CCAATGAGCGCTGCTGTGCTAAATTTTGTGTAATCACCACAAAGATGGGGCTGGAATG 749

QY 572 ATGTTATCTGTAAGACTAGAAGGAATTCATATGTGAGATGAATAATTTACCTA 627
DB 750 ATGTTAATTTGCTGGTCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACTTA 805

RESULT 15
AX400064 1195 bp DNA linear PAT 06-JUN-2002
LOCUS Sequence 235 from Patent W00218424.
DEFINITION AX400064
ACCESSION AX400064
VERSION AX400064.1 GI:21336479
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Tang, Y.T., Asundi, V., Zhou, P., Xue, A.J., Ren, F., Zhang, J.,
Wang, J.R., Zhao, Q.A., Wang, D., Liu, C., Drmanac, R.T. and Wehrman, T.
Nucleic acids and polypeptides
Patent: WO 0218424-A 235 07-MAR-2002;
JOURNAL HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
source 1. .1195
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 270. .866
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD33508.1"
/db_xref="GI:21336480"
/translation="MTSEITYAEVRFKNEFKSSGINTASSAVFFOKYSOLLEKKTKE
LVHTTLECYKKNPVRETAWSCCPKNKWSFSSNCYFFISTESASWQSEKDCARMEAHLL
LVINTQEEODFIQNLOEESAYFVGLSDPEGRHQWQVDQTPYNESSTFWHPREPSDP
NERCVVLNFRKSPKRWGWNVNDVCLGPQRSVCVMKHL"
```

BASE COUNT		340	a	224	c	268	g	363	t
ORIGIN									
Query Match		30.6%; Score 192; DB 6; Length 1195;							
Best Local Similarity		68.0%; Pred. No. 5.5e-46;							
Matches		283;	Conservative	0;	Mismatches	130;	Indels	3;	Gaps 1;
QY	215	AGGTGCCAGCCTGGGGATGTTGCCACAGCTTCTTGGAAATCAATTTGGTTCCAGTTGCTACT	274						
DB	448	AAGAGACAGCCTGGAGCTGTTGCCCAAGAATTTGGAAGTCATTTAGTTCCAACTGCTACT	507						
QY	275	TCATTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGTGCAGCAACTCTGTTGAGATGGGAG	334						
DB	508	TTATTTCTACTGAATCAGCATCTTGGCAACACAGTGAGAAGACTGTGCTAGAAITGGAGG	567						
QY	335	CACATTTGGTTGTGTTCAACACAGACAGCAGAGAAATTTCAATTTGCCAGCAGCTGAATG	394						
DB	568	CTCACCTGCTGGTGATAAACACTCAAGAAGAGCAGGATTTTCATCTTCCAGAATCTGCAAG	627						
QY	395	AGTCATTTTCTTATTTTCTGGGGCTTTTCAGACCCACAGGTAATAATATTTGGCAATGGA	454						
DB	628	AAGAATCTGCTTATTTTGTGGGGCTCTCAGATCCAGAGGTCAGCGACATTTGGCAATGGG	687						
QY	455	TTGATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCATT	514						
DB	688	TTGATCAGACACCATACATGAAGTTCCACATTTCTGGCATCCACGTGAGCCCACTGATC	747						
QY	515	CTCAGAGCAATGTGCTTCAATAGTCTTCTCGAA---ACCTACAGGATGGGGCTGGAATG	571						
DB	748	CCAATGAGCGCTGCGTTGTGCTAAATTTTCGTAAATCACCCAAAGATGGGGCTGGAATG	807						
QY	572	ATGTTATCTGTGAAGTGAAGAAATTCATATGTGAGATGAATTAAGATTTACCTA	627						
DB	808	ATGTTAATTGCTCTTGGTCTCCTCAAAGGTCAGTTTGTGAGATGATGAAGATCCACTTA	863						

Search completed: February 17, 2003, 11:03:47  
Job time : 1085.37 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 20, 2003, 19:28:17 ; Search time 47 Seconds  
(without alignments)  
1363.733 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09766511/runat\_19022003\_094339\_29316/app\_query.fasta\_1.391  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1\_1\_17.erunat\_19022003\_094339\_29316 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	793	68.1	1227	3	US-08-772-440-3
2	679	58.3	501	3	Sequence 3, Appli
3	565	48.5	393	3	Sequence 20, Appl
4	458	39.3	1104	4	Sequence 22, Appl
5	436	37.4	1418	4	Sequence 1, Appli
6	251.5	21.6	1458	4	Sequence 7, Appli
7	250	21.5	1370	4	Sequence 3, Appli
8	235	20.2	1212	4	Sequence 9, Appli
9	235	20.2	1212	4	Sequence 1, Appli
10	234	20.1	1212	4	Sequence 11, Appl
11	232	19.9	1212	4	Sequence 10, Appl
12	229.5	19.7	10409	3	Sequence 33, Appl

13	223	19.1	152	3	US-08-772-440-40	Sequence 40, Appl
14	192.5	16.5	4771	3	US-08-840-062-3	Sequence 3, Appli
15	191	16.4	145	3	US-08-772-440-42	Sequence 42, Appl
16	189.5	16.3	4588	3	US-08-840-062-1	Sequence 1, Appli
17	187	16.1	1897	2	US-08-809-494A-1	Sequence 1, Appli
18	187	16.1	1897	4	US-09-352-302-1	Sequence 1, Appli
19	182.5	15.7	3259	5	PCT-US95-03747-1	Sequence 3, Appli
20	180.5	15.5	885	1	US-08-365-103B-3	Sequence 1, Appli
21	180.5	15.5	924	1	US-08-365-103B-5	Sequence 5, Appli
22	180.5	15.5	1005	1	US-08-365-103B-1	Sequence 1, Appli
23	179	15.4	1306	2	US-08-809-494A-3	Sequence 3, Appli
24	179	15.4	1306	4	US-09-352-302-3	Sequence 3, Appli
25	172	14.8	116	3	US-08-772-440-41	Sequence 41, Appl
26	169.5	14.5	561	4	US-09-535-521-16	Sequence 16, Appl
27	169.5	14.5	561	4	US-09-535-521-18	Sequence 18, Appl
28	169.5	14.5	624	4	US-09-535-521-19	Sequence 19, Appl
29	169.5	14.5	624	4	US-09-535-521-21	Sequence 21, Appl
30	169.5	14.5	876	4	US-09-535-521-4	Sequence 4, Appli
31	169.5	14.5	876	4	US-09-535-521-6	Sequence 6, Appli
32	169.5	14.5	2851	4	US-09-535-521-1	Sequence 1, Appli
33	169.5	14.5	2851	4	US-09-535-521-3	Sequence 3, Appli
34	168	14.4	996	4	US-09-127-946-11	Sequence 11, Appl
35	167.5	14.4	1461	3	US-08-722-126A-4	Sequence 4, Appli
36	167.5	14.4	1461	5	PCT-US95-04258-4	Sequence 10, Appl
37	166.5	14.3	417	4	US-09-535-521-10	Sequence 12, Appl
38	166.5	14.3	417	4	US-09-535-521-12	Sequence 13, Appl
39	166.5	14.3	423	4	US-09-535-521-13	Sequence 15, Appl
40	166.5	14.3	423	4	US-09-535-521-15	Sequence 7, Appli
41	166	14.2	384	4	US-09-535-521-7	Sequence 9, Appli
42	166	14.2	384	4	US-09-535-521-9	Sequence 2, Appli
43	166	14.2	5169	4	US-09-194-612A-2	Sequence 24, Appl
44	165.5	14.2	369	4	US-09-535-521-24	Sequence 26, Appl
45	165.5	14.2	369	4	US-09-535-521-26	

ALIGNMENTS

RESULT 1  
US-08-772-440-3  
; Sequence 3, Application US/08772440  
; Patent No. 6046158  
; GENERAL INFORMATION:  
; APPLICANT: Arizumi, Kiyoshi  
; APPLICANT: Takashima, Akira  
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE  
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/772.440  
; FILING DATE: CONCURRENTLY HERewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UTXD.493  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:

1



```
Db 481 GAAATGAAGAAGATTACCTA 501
RESULT 3
US-08-772-440-22
; Sequence 22, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Arizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 22:
; LENGTH: 393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-22
Alignment Scores:
Pred. No.: 1,52e-57 Length: 393
Score: 565.00 Matches: 96
Percent Similarity: 81.68% Conservative: 11
Best Local Similarity: 73.28% Mismatches: 24
Query Match: 48.50% Indels: 0
DB: 3 Gaps: 0
US-09-766-511b-53 (1-209) x US-08-772-440-22 (1-393)
Qy 79 CysProAlaSerTrpLysSerPheGlySerCysTyrPheLeuSerGluGluLys 98
|||||
Db 1 TGCCCAATCACTGGAAGTCATTGGCTCCAGCTGCTACCTCATCTTACCAAGGAGAAC 60
Qy 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsn 118
|||||
Db 61 TTCCTGGAGCACCGACGAGCACTGTTCAGATGGGGCTCATCTGGTGGTATCAAT 120
Qy 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTrpPheLeu 138
|||||
Db 121 ACTGAACGGAGAGAGAAATTCATCACCACAGCTGAATGAGTCACCTTCTTACTTCCTG 180
Qy 139 GlyLeuSerAspProGlnGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGlu 158
|||||
Db 181 GGTCTTCGGATCCACCAAGGTAAATGGCAATGGCAATGGATCGATGATCTCTCTTCAGT 240
Qy 159 LysAsnValArgPheTrpHisLeuGlyCluProAsnHisSerAlaGluGlnCysAlaSer 178
|||||
Db 241 CAAAATGTCAGGTCTTGGCAGCCCATGAACCCCAATCTTCCAGAAGACGGTGTGTTTCA 300
Qy 179 IleValPheTrpLysProThrGlyTrpAsnAspValIleCysGluThrArgArg 198
|||||
Db 301 ATAGTTTACTGGAATCCCTTCGAAATGGGCTGGATGATGTTTCTGTGATAGTAACAC 360
Qy 199 AsnSerIleCysGluMetAsnLysIleTyrLeu 209
|||||
Db 361 AATCAATATGTGAATGAAGAAGATTACCTA 393
RESULT 4
US-09-111-470-1
; Sequence 1, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 242..952
US-09-111-470-1
Alignment Scores:
Pred. No.: 2,82e-44 Length: 1104
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 4 Gaps: 3
US-09-766-511b-53 (1-209) x US-09-111-470-1 (1-1104)
Qy 26 TieSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValIleThrHisPhe 45
|||||
```



Db 954 AACAGAGTCAGTTTCTCAGATGAAGAAAAATAACTTA 992

RESULT 6

US-09-111-470-3

; Sequence 3, Application US/09111470

; Patent No. 6277959

; GENERAL INFORMATION:

; APPLICANT: Valladeau, Jenny

; APPLICANT: Ravel, Odile

; APPLICANT: Bates, Elizabeth E.M.

; APPLICANT: Ford, John

; APPLICANT: Saeland, Sem

; APPLICANT: Lebecque, Serge J.E.

; TITLE OF INVENTION: Mammalian Membrane Protein Genes;

; TITLE OF INVENTION: Related Reagents

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DNAX Research Institute

; STREET: 901 California Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/111,470

; FILING DATE: 08-JUL-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/053,080

; FILING DATE: 09-JUL-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ching, Edwin P.

; REGISTRATION NUMBER: 34,090

; REFERENCE/DOCKET NUMBER: SF0695

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650)852-9196

; TELEFAX: (650)496-1200

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1458 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 257..1204

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 608

; OTHER INFORMATION: /note= "short form lacks

; OTHER INFORMATION: nucleotides 608-673"

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 775

; OTHER INFORMATION: /note= "ASGPRm (table 2) has

; OTHER INFORMATION: sequence insert encoding GEE between nucleotides 775-776"

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 1064

; OTHER INFORMATION: /note= "nucleotide 1064 of DCMp2s

; OTHER INFORMATION: may be A, which would encode asn rather than asp at the resid

; OTHER INFORMATION: numbered 270"

US-09-111-470-3

Alignment Scores:

Pred. No.:

Score:

7.66e-20

251.50

Length:

Matches:

1458

55

Percent Similarity: 49.70% Conservative: 28  
Best Local Similarity: 32.93% Mismatches: 55  
Query Match: 21.59% Indels: 29  
DB: 4 Gaps: 7

US-09-766-511b-53 (1-209) x US-09-111-470-3 (1-1458)

QY 53 ArgLeuSerGluLeuHisSerTyrHisSerSerLeuThrCys----- 66

Db 710 CGAGTCCAGCAGCTGGTGCAGAACCTGGAAGAACTGACCTGCCAGGTGCTACTCTCAAC 769

QY 67 -----PheSerGluGlyThrLysValProAlaTrpGlyCysCysProAlaSerTrp 83

Db 770 AACAAATGCTCCAGTCAAGGACC-----TGTGCCCCCTCAACTGG 811

QY 84 LysSerPheGlySerSerCysTyrPheLeSerSerGluGluLysValTrpSerLysSer 103

Db 812 GTGGAGCACCAAGACAGCTGCTACTGTTCTCTCCTCTGGGATGCTCTGGCGGAGGCT 871

QY 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123

Db 872 GAGAACTACTGCCAGCTGAAGAACGCCACCTGGTGGTCACTCACTCCAGGAGGAGGAG 931

QY 124 AsnPheIleValGlnGlnGluAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143

Db 932 AATTTTGTCCAGAAATATCTAGGCTCCGATACACCTGG---ATGGCCTCAGTCACCT 988

QY 144 GlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163

Db 989 GAAGGA-----GCCTGGAAGTGGGTGGATGGAACAGACTATATCGACCGGCTTCCAGA 1042

QY 164 TrpHisLeuGlyGluProAsnHis-----SerAlaGluGlnCys 176

Db 1043 TGAAGCCAGGCCAGCCAGCAGCTGGAGGGGACGGGCTGGGTGGAGGGCAGACTGT 1102

QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpGlyTrpGlyTrpGlyTrpGlyTrp 196

Db 1103 GCT-----CACTTCCATCCAGACGGC---AGGTGGAATGACGACGCTGTGCCAGAG 1150

QY 197 ArgArgAsnSerIleCysGlu 203

Db 1151 CCTACCACCTGGGTCTGCGAG 1171

RESULT 7

US-09-111-470-9

; Sequence 9, Application US/09111470

; Patent No. 6277959

; GENERAL INFORMATION:

; APPLICANT: Valladeau, Jenny

; APPLICANT: Ravel, Odile

; APPLICANT: Bates, Elizabeth E.M.

; APPLICANT: Ford, John

; APPLICANT: Saeland, Sem

; APPLICANT: Lebecque, Serge J.E.

; TITLE OF INVENTION: Mammalian Membrane Protein Genes;

; TITLE OF INVENTION: Related Reagents

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DNAX Research Institute

; STREET: 901 California Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/111,470

; FILING DATE: 08-JUL-1998

; CLASSIFICATION:

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 273..1091
US-09-111-470-9

Alignment Scores:
Pred. No.: 1,04e-19 Length: 1370
Score: 250.00 Matches: 55
Percent Similarity: 48.82% Conservative: 28
Best Local Similarity: 32.35% Mismatches: 55
Query Match: 21.46% Indels: 32
DB: 4 Gaps: 7

US-09-766-511B-53 (1-209) x US-09-111-470-9 (1-1370)
QY 53 ArgLeuSerGluHisSerTyrHisSerSerLeuThrCys----- 66
   |||...:|||||
Db 588 CGAGTCACAGCAGCTGGTCAGACAGCTGAGAACTGACCTGCCAGGTGGTACTCTCAAC 647
QY 67 -----PheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
   :|||||
Db 648 AACAATGGTGAGGAAGCCTCCACTCAAGGACC-----TGCTGCCCC 689
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluLysValTrp 100
   :|||||
Db 690 GTCAACTGGGTGGACACACAGCAGCTGCTACTGGTTCTCTCACTCTGGGATGCTCTGG 749
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
   :|||||
Db 750 GCCAGGCTGAGAGTACTGCCAGCTGAGAACGCCACCTGGTGGTCATCACTCCAGG 809
QY 121 AlaGluGlnAsnPheIleValGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
   :|||||
Db 810 GAGGACAGCAATTTGTCAGAAATATCTAGGCTCCGCATACAGCTGG---ATGGGCTC 866
QY 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
   :|||||
Db 867 AGTGACCTGTAAGGA-----GCCTGGAAGTGGGTGGATGAACAGACTATGCCACCGGC 920
QY 161 ValArgPheTrpHisLeuGlyGluProAsnHis-----SerAla 173
   :|||
Db 921 TTCAGAACTGGAAGCCAGCCAGCCAGCAGCAGCTGGCAGGGCCAGCGGCTGGGTGGAGGC 980
QY 174 GluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIle 193
   |||
Db 981 GAGGACTGTGCT-----CACTTCCATCCAGACGCGC---AGTGGAAATGACGACGCTC 1028
QY 194 CysGluThrArgArgAsnSerIleCysGlu 203
   |||
Db 1029 TCCAGAGGCCCTACCACCTGGGTCTGCCAG 1058

RESULT 8
US-09-591-435-9
; Sequence 9, Application US/09591435
; Patent No. 6280953
; GENERAL INFORMATION:
; APPLICANT: Messier, Walter
; APPLICANT: SIKELA, JAMES M
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
; TITLE OF INVENTION: AND MEDICAL CONDITIONS
; FILE REFERENCE: GENE.200.2
; CURRENT APPLICATION NUMBER: US/09/591,435
; PRIOR APPLICATION NUMBER: 09/591,435
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/240,915
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/073,263
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/098,987
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-591-435-9

Alignment Scores:
Pred. No.: 4.95e-18 Length: 1212
Score: 235.00 Matches: 48
Percent Similarity: 55.47% Conservative: 23
Best Local Similarity: 37.50% Mismatches: 49
Query Match: 20.17% Indels: 8
DB: 4 Gaps: 4

US-09-766-511B-53 (1-209) x US-09-591-435-9 (1-1212)
QY 79 CysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLys 98
   |||...:|||||
Db 766 TGTCCTCGGGAATGGACATCTTCCAGGAAGACTGTACTTCTAATGCTCACTCCAGCGG 825
QY 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118
   |||
Db 826 AACTGGCAGCAGCTCCATCATCCGCGCTGCAAGAAGAGTGGGGCCAGCTCGCTGAATCAAA 885
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr---Phe 137
   :|||
Db 886 AGTCTGAGGAGAGCAGAACTCTCTACAGCTGCAGTCTTCCAGAAAGTAACCGCTTCACTCG 945
QY 138 LeuGlyLeuSerAspProGlnGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
   :|||
Db 946 ATGGGACTTTCAGATCTAAATCAGGAAGGCACGCTGGCAATGGGTGGCAGGCTCACCTCTG 1005
QY 158 GluLysAsnValArg---PheTrpHisLeuGlyGluProAsnHisSerAlaGluGln--- 175
   :|||
Db 1006 TTGCCAGCTTCAAGCAGTATTGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
QY 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195
   |||
Db 1066 TGGCGCGGAATTT-----AGTGGCAATGGCTGGCAACGACGACCAATGAATTAAT 1110
QY 196 ThrArgArgAsnSerIleCysGlu 203
   :|||
Db 1111 CTTGCCAAATCTGGATCTGCAAA 1134

RESULT 9
US-09-517-605-1
; Sequence 1, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
```





```
Db 5935 TAGGAAATGTAGTTGAAATGTATGTTTCATCCCATTTAGAAAAGCCCTTAGAATTTT 5994
QY 137 e-----LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAs 153
      |||
Db 5995 TATACCTAGGAGATTGGAGTTCTCTCCTCAGTGGTAGTCGAAT----- 6040
QY 153 pLysThrProTyrGluLysAsnValArgPheTrpHisLeu 166
      ::|||
Db 6041 -AGACTACCCACTTCCCAATCCATAAGGTTCTGGCAATTA 6079

RESULT 13
US-08-772-440-40
; Sequence 40, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/772,440
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-40

Alignment Scores:
Pred. No.: 4,82e-18 Length: 152
Score: 223.00 Matches: 38
Percent Similarity: 85.42% Conservative: 3
Best Local Similarity: 79.17% Mismatches: 7
Query Match: 19.14% Indels: 0
DB: Gaps: 0

US-09-766-511b-53 (1-209) x US-08-772-440-40 (1-152)
QY 76 TrpGlyCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSer 95
      |||
Db 9 TGGGGATGCTGCCCAATCACTGGAGTCATTTGGCTCCAGCTGCATCTTCTACC 68
      ::|||
QY 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
      ::|||
Db 69 AAGGAGAACTTCTGGAGCACCAGTGAGCAGAACTGTGTTCCAGATGGGGCTCATCTGGTG 128
      |||
QY 116 ValPheAsnThrGluAlaGluGln 123
      |||
Db 129 GTGATCAATACTAAGCGGAGCAG 152
```

```
RESULT 14
US-08-840-062-3
; Sequence 3, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: WU, KAI
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/840,062
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019r1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-840-062-3

Alignment Scores:
Pred. No.: 4,19e-12 Length: 4771
Score: 192.50 Matches: 43
Percent Similarity: 52.71% Conservative: 25
Best Local Similarity: 33.33% Mismatches: 52
Query Match: 16.52% Indels: 9
DB: Gaps: 5

US-09-766-511b-53 (1-209) x US-08-840-062-3 (1-4771)
QY 79 CysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLys 98
      |||
Db 1156 TCGAGCGCGAGCTGGCGAGCCCTTCCAGGGCCACTGCTACCGCTGCAGGCCGAGAAGCGC 1215
      |||
QY 99 ValTrpSerLysSerClnAsnCysValGluMetGlyAlaHisLeuValPheAsn 118
      |||
Db 1216 AGCTGGCAGGAGTCCAAAGAGGCATGCTACGGGGGGTGGCGACCTGGTCAGCATCCAC 1275
      |||
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPhe---SerTyrPhe 137
      ::|||
Db 1276 AGCATGGCGGAGCTGGAATTCATCACCAGCAGATCAACGAGAGGTGGAGGAGCTGTGG 1335
      |||
QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
      ::|||
Db 1336 ATCGGCTCAACGATTTGAAGCTGCAGATGAATTTTGTGTCTCAGCGGGAGCCTTGTG 1395
      |||
QY 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHis-----SerAlaGlu 174
      ::|||
Db 1396 -----AGCTTCACCCACTGGGACCCCTTTGAGCCCAACAACTCCGGGACAGTCTGGAG 1449
      |||
QY 175 GlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCys 194
```

Job time : 52 secs

```
Db 1450 GACTGTGTACCATC-----TGGGGCCCGGAAGGC---CGCTGGAGCAGACAGTCCCTGT 1500
QY 195 GluThrArgAsnSerIleCysGlu 203
Db 1501 AACCACTCTTGCATCCATCTGCAAG 1527

RESULT 15
US-08-772-440-42
; Sequence 42, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Arizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-42

Alignment Scores:
Pred. No.: 2.56e-14 Length: 145
Score: 191.00 Matches: 30
Percent Similarity: 76.60% Conservatives: 6
Best Local Similarity: 63.83% Mismatches: 11
Query Match: 16.39% Indels: 0
DB: 3 Gaps: 0

US-09-766-511B-53 (1-209) x US-08-772-440-42 (1-145)
QY 163 PheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrp 182
Db 2 TTCTGGCACCCCATGATGAACCCCAATCTCCAGAGAGGGGTGTTCATAGTTTACTGG 61
QY 183 LysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
Db 62 AATCCTTCGAATGGGGCTGGAATGATGTTTCTGTGATAGTAAACACAATTCATATGT 121
QY 203 GluMetAsnLysIleTyrLeu 209
Db 122 GAAATGAAGAAGATTACCTA 142

Search completed: February 20, 2003, 20:40:40
```



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 18:36:42 ; Search time 220 Seconds  
(without alignments)  
2139.398 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPOSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Listing first 45 summaries

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094338\_29278/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -FRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pto -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1\_125\_@runat\_19022003\_094338\_29278 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1165	100.0	3114	22	AAS01375	Human TANGO 405 cd
2	892.5	76.6	850	20	AAZ07531	Human SDCMP3 polyp
3	793	68.1	630	20	AAZ07532	Mouse SDCMP3 polyp
4	793	68.1	1227	19	AAV42549	Mouse dectin-2 cdn
5	793	68.1	1252	22	AAS01378	Murine TANGO 405 a
6	740	63.5	821	22	AAS01376	Murine TANGO 405 c
7	679	58.3	501	19	AAV42554	Mouse dectin-2 ext
8	637	54.7	1312	22	AAF90241	Nucleotide sequenc
9	630	54.1	827	22	AAF19729	Dendritic cell (DC
10	617	53.0	800	22	AAJ19730	Dendritic cell (DC
11	459.5	39.4	444	24	ABK52904	Human cdna encodin
12	459	39.4	402	24	ABK52903	Human cdna encodin
13	458	39.3	1091	22	AAS31385	Human cdna encodin
14	458	39.3	1091	24	ABO66709	Human polynucleoti
15	458	39.3	1096	22	AAS31224	Human cdna encodin
16	458	39.3	1096	24	ABO66548	Human polynucleoti
17	458	39.3	1104	20	AAK04865	Primate DCMPI C-le
c 18	458	39.3	1307	24	ABL90686	Human polynucleoti
c 19	439.5	37.7	582	24	ABK52904	Human dendritic ce
c 20	436	37.4	1418	20	AAK04867	rodent DCMPI C-lec
c 21	427.5	36.7	568	24	ABK52908	Human dendritic ce
c 22	419	36.0	558	24	ABK52907	Human dendritic ce
c 23	400.5	34.4	403	24	ABK52909	Human dendritic ce
c 24	398.5	34.2	968	22	AAF81743	Human membrane ass
c 25	398.5	34.2	997	20	AAK52274	Protein PRO244 cdn
c 26	398.5	34.2	997	22	AAK52936	Human DNA encoding
c 27	398.5	34.2	997	22	AAF72432	Human PRO244 cdna
c 28	390.5	33.5	2517	22	AAF66349	Murine cdna encodi
c 29	382	32.8	1013	24	ABK52905	Human dendritic ce
c 30	377.5	32.4	753	24	ABK52906	Human dendritic ce
c 31	375.5	32.4	561	24	ABK52911	Human dendritic ce
c 32	356.5	30.6	2059	21	AAA26386	Human secreted pro
c 33	356	30.6	324	22	AAK63162	Human immune/haema
c 34	303.5	26.1	323	22	AAK53734	Murine transport a
c 35	299.5	25.7	448	24	ABK52912	Human dendritic ce
c 36	290.5	24.9	666	24	ABK52910	Human dendritic ce
c 37	267.5	23.0	559	24	ABK52913	Human dendritic ce
c 38	254	21.8	1290	24	ABK63735	Rat sequence diffe
c 39	252.5	21.7	1309	24	ABN95737	Gene #2235 used to
c 40	252	21.6	1300	18	AAK66951	Asialoglycoprotein
c 41	252	21.6	1300	24	ABN97236	Gene #3734 used to
c 42	251.5	21.6	1458	20	AAK04866	Primate DCMPI C-le
c 43	250	21.5	1370	20	AAK04868	Variant primate DC
c 44	250	21.5	1531	24	ABK94927	Human novel polynu
c 45	249	21.4	2005	24	ABL95574	Human angiogenesis

ALIGNMENTS

RESULT 1  
AAS01375  
ID AAS01375 standard; cdna: 3114 BP.  
XX  
XX AAS01375;  
AC  
AC  
DT 04-JUL-2001 (first entry)  
XX  
DE Human TANGO 405 cdna sequence.  
KW Human; TANGO 210; clone jthLai52H06; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.  
XX  
OS Homo sapiens.  
XX

PH	Key	Location/Qualifiers	
FT	CDS	154..783	
FT		/*tag= a	
FT		/product= "TANGO 405 protein"	
FT		/note= "The ORF is specifically claimed"	
FT	sig_peptide	154..297	
FT		/*tag= b	
FT	mat_peptide	298..780	
FT		/*tag= c	
XX			
PN	WO200118016-A1.		
XX			
PD	15-MAR-2001.		
XX			
PF	30-JUN-2000; 2000WO-US18174.		
PR	10-SEP-1999; 99US-0393996.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;		
DR	WPI; 2001-183280/18.		
XX	P-PSDB; AAU00479.		
XX			
PT	Isolated nucleic acid molecules encoding proteins useful as modulating		
PT	agents in regulating a variety of cellular processes are used for		
PT	treating e.g. cancer and autoimmune disorders -		
XX			
PS	Claim 2; Fig 6A-6C; 326pp; English.		
XX			
CC	The present sequence encoding for human TANGO 405 protein is isolated		
CC	from cDNA clone jthai152h06 from a human mixed lymphocyte reaction cDNA		
CC	library. It is 1 of 6 novel human proteins which include TANGO 210		
CC	(AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394		
CC	(AAU00473) and INTERCEPT 400 (AAU00476). Novel sequences for murine		
CC	TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and		
CC	a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic		
CC	acids encoding these novel proteins are useful as modulating agents in		
CC	regulating a variety of cellular processes and can be used to express		
CC	the proteins in a host cell in gene therapy applications. Human and		
CC	murine TANGO 405 proteins show sequence homology to murine dectin-2.		
CC	TANGO 405 modulates growth, proliferation, survival, differentiation,		
CC	activity, morphology and movement/migration of human lymphocytes and		
CC	bone marrow cells and tissues and can be used to prevent, diagnose or		
CC	treat leukaemia, lymphomas and autoimmune disorders.		
XX			
SQ	Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 Other;		
Alignment Scores:			
Pred. No.:	4e-116	Length:	3114
Score:	1165.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0
US-09-766-511b-53 (1-209) x AAS01375 (1-3114)			
QY	1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20		
DB	154 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGGCTGGTGTCCCTGAGACTC 213		
QY	21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40		
DB	214 TGGTCTGTGGCTGGGATTTCATTTGCACCTCTCAGTGTGCTTCATTTGAGCTGTGTA 273		
QY	41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60		
DB	274 GTAACCTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACATCATAT 333		
QY	61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80		

Db

334

CATTCAAGTCTCACCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA

393

QY

81

AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluLysValTrp

100

Db

394

GCITCTTGGGAAGTCATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAGGTTTGG

453

QY

101

SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu

120

Db

454

TCTAAGAGTCAGCAGAACTGTGTGAGATGGGAGCACATTTGGTTGTTCACACAGAA

513

QY

121

AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu

140

Db

514

GCAGACAGAAATTCATTTGCCAGCAGCTGAATGAGTCATTTTCTTATTTTCTGGGGCTT

573

QY

141

SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn

160

Db

574

TCAGACCCACAAAGTAAATAATTTGGCAATGGATTGATGAAGACACCTTATGAGAAAAAT

633

QY

161

ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal

180

Db

634

GTCAAGATTTTGGCACCCTAGGTGAGCCCAATCATTTCTGCAGACCAATGTCTTCAATAGTC

693

QY

181

PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer

200

Db

694

TTCTGGAACCTACAGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAAGGAATTCA

753

QY

201

IleCysGluMetAsnLysIleTyrLeu

209

Db

754

ATATGTGAGATGAATAAGATTTACCTA

780

RESULT 2

AAZ07531

ID

AAZ07531 standard; cDNA; 850 BP.

XX

AAZ07531;

XX

26-NOV-1999 (first entry)

XX

Human SDCMP3 polypeptide encoding cDNA.

XX

Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;

XX

Schering dendritic cell membrane protein; dendritic cell physiology;

XX

genetic fingerprinting; cancer immunotherapy; abnormal proliferation;

XX

cancer; forensic; human; lectin 73; ss.

OS

Homo sapiens.

XX

Key

Location/Qualifiers

FT

CDS

108..596

FT

/\*tag= a

FT

/product= "SDCMP3"

XX

WO9947673-A2.

XX

23-SEP-1999.

XX

16-MAR-1999; 99WO-US03740.

XX

17-MAR-1998; 98US-0040111.

XX

(SCHE ) SCHERING CORP.

XX

Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;

XX

Philipps JH;

XX

WPI; 1999-562114/47.

DR

P-PSDB; AAY27446.

XX

Binding compound specific for primate or rodent Schering dendritic cell

XX

membrane proteins -

XX

Claim 8; Page 80; 89pp; English.

XX

CC The invention relates to a binding compound comprising an antibody  
 CC binding site which specifically binds to: (a) a primate or rodent  
 CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
 CC primate SDCMP4 protein. The binding compound can be used to modulate  
 CC dendritic cell physiology or function. The sequences may also be used as  
 CC probes in forensic techniques, such as genetic fingerprinting. They can  
 CC also be used to distinguish tissue and cell types in situ or in vitro.  
 CC The sequences may also be useful in cancer immunotherapy and in the  
 CC treatment of conditions associated with abnormal physiology or  
 CC development, including abnormal proliferation, e.g. cancerous conditions,  
 CC or degenerative conditions. The present sequence represents a cDNA  
 CC encoding the human SDCMP3 (lectin 73).  
 XX

SQ Sequence 850 BP; 257 A; 153 C; 188 G; 252 T; 0 other;

#### Alignment Scores:

Pred. No.: 2,54e-87 Length: 850  
 Score: 892.50 Matches: 165  
 Percent Similarity: 97.63% Conservative: 0  
 Best Local Similarity: 97.63% Mismatches: 1  
 Query Match: 76.61% Indels: 3  
 DB: 20 Gaps: 1

US-09-766-511B-53 (1-209) x AAZ07531 (1-850)

QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20  
 DB 108 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGCTGGTTCCTCGAGACTC 167  
 QY 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
 DB 168 TGGTCTGTGGCTGGATTTCCATTGCACCTCTCAGTGTCTTCATTGTGAGCTGTGTA 227  
 QY 41 ValThrTrpHisPheThrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTrp 60  
 DB 228 GTACTTACCATTTCATATGTTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATAT 287  
 QY 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
 DB 288 CATTCAAGTCTTACCTGCTTCAAGGAGCAAGGTGCACGCTGGGATGTGCCCA 347  
 QY 81 AlaSerTrpLysSerPheGlySerSerCysTrpPheIleSerSerGluGluLysValTrp 100  
 DB 348 GCITTCCTGGAGTCATTGGTTCAGTTCGCTACTTTCATTCCAGTGAAGAGAGGTTTG 407  
 QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
 DB 408 TCTAAGAGTCAGCAGAACTGTGTGAGATGGGAGCACATTGGTTGTTCACACAGAA 467  
 QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTrpPheLeuGlyLeu 140  
 DB 468 GCAGAGCAGAAATTCATTGTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGCTT 527  
 QY 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTrpGluLysAsn 160  
 DB 528 TCAGACCCACAGGTAAATAATTTGGCAATGGATTTGATAAGACACCTTATGAGAAAAAT 587  
 QY 161 ValArg-----PheTrpHisLeu 166  
 DB 588 GTCAGGTGAGTCAGTCTCTGGGCGCTT 614

#### RESULT 3

AAZ07532  
 ID AAZ07532 standard; cDNA; 630 BP.

XX

AC AAZ07532;

XX 26-NOV-1999 (first entry)

DE Mouse SDCMP3 polypeptide encoding cDNA.

XX Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;

KW Schering dendritic cell membrane protein; dendritic cell physiology;  
 KW

KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
 KW cancer; forensic; mouse; ss.

XX Mus sp.

XX Key Location/Qualifiers  
 FH 1..630  
 CDS /\*tag= a  
 FT /product= "SDCMP3"

XX W09947673-A2.

XX 23-SEP-1999.

XX 16-MAR-1999; 99WO-US03740.

XX 17-MAR-1998; 98US-0040111.

XX (SCHE ) SCHERING CORP.

PI Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
 PI Philipps JH;

DR WPI; 1999-562114/47.  
 DR P-PSDB; AAV27447.

PT Binding compound specific for primate or rodent Schering dendritic cell  
 PT membrane proteins

PS Claim 8; Page 82; 89pp; English.

CC The invention relates to a binding compound comprising an antibody  
 CC binding site which specifically binds to: (a) a primate or rodent  
 CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
 CC primate SDCMP4 protein. The binding compound can be used to modulate  
 CC dendritic cell physiology or function. The sequences may also be used as  
 CC probes in forensic techniques, such as genetic fingerprinting. They can  
 CC also be used to distinguish tissue and cell types in situ or in vitro.  
 CC The sequences may also be useful in cancer immunotherapy and in the  
 CC treatment of conditions associated with abnormal physiology or  
 CC development, including abnormal proliferation, e.g. cancerous conditions,  
 CC or degenerative conditions. The present sequence represents a cDNA  
 CC encoding the mouse SDCMP3.

SQ Sequence 630 BP; 171 A; 135 C; 152 G; 172 T; 0 other;

#### Alignment Scores:

Pred. No.: 1.07e-76 Length: 630  
 Score: 793.00 Matches: 145  
 Percent Similarity: 77.00% Conservative: 19  
 Best Local Similarity: 68.08% Mismatches: 41  
 Query Match: 68.07% Indels: 8  
 DB: 20 Gaps: 4

US-09-766-511B-53 (1-209) x AAZ07532 (1-630)

QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
 DB 1 ATGGTGCAGGAAGACAAATCCAGGG-----AAGGAGTCTGCTGG---ACCTGT 48  
 QY 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
 DB 49 AGACTCTGTCAGCTGCTGTGATTTCCATGTTACTCTTGTAGTACCTGTTTCTTCGAGC 108  
 QY 39 CysValValThrTrpHisPheThrTrpGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
 DB 109 TGTGTGGTGACTTACCAATTTATTATGACACAGCCAGCCAGTATATGAACCTTCC 168  
 QY 59 SerTrpHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
 DB 169 ACATACCATTTCAGTCTCAGCTGCTTCAGTGAAGGAGACTATGGTGTCTCAGAAAAATGTG 228  
 QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTrpPheIleSerSerGlu 96

Db 229 GGATGCTGCCAAATCACTGGAAGTCATTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 288  
 QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
 Db 289 GAGAACTTCTGGAGCACCAGTGCAGAGAACCTGTTCAGATGGGGCTCATCTGGTGGTG 348  
 QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
 Db 349 ATCAATACTGAAGCGGAGCAGAAATTCATACCAGCAGCTGAATGAGTCACTTCTTAC 408  
 QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 Db 409 TTCTTGGTCTTTCGGATCCACAGGTAAATGGCAATGGCAATGGATGATGATCTCCT 468  
 QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 Db 469 TTCAGTCAAAATGTCAGTTCTGGCACCCCATGACCCCAATCTCCAGAGAGCGGTGT 528  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 Db 529 GTTTCATAGTTTACTGGAATCCTTCGAAATGGGGCTGGAATGATGTTTCTGTGATAGT 588  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 589 AAACACAATTCATATGTGAAATGAAGAAGATTACCTA 627  
 RESULT 4  
 AAV42549  
 ID AAV42549 standard; cDNA; 1227 BP.  
 AC AAV42549;  
 DT 09-NOV-1998 (first entry)  
 DE Mouse dectin-2 cDNA.  
 KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
 KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
 KW drug screening; ss.  
 XX  
 OS Mus sp.  
 FH Key  
 FT CDS  
 FT 146..775  
 FT /\*tag= a  
 XX  
 PN W09828332-A2.  
 PD 02-JUL-1998.  
 PF 22-DEC-1997; 97WO-US23761.  
 PR 20-DEC-1996; 96US-0772440.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA Ariizumi K, Takashima A;  
 PI WPI: 1998-377594/32.  
 DR P-PSDB; AAW63010.  
 XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 PT and -2 - useful, e.g. to regulate immune response, as vaccine  
 PT adjuvants, for diagnosis and drug screening  
 XX  
 PS Claim 43; Page 141-142; 200pp; English.  
 XX This cDNA clone codes for a novel murine dendritic cell (DC)  
 CC specific member of the C-type lectin family, termed DC-associated  
 CC C-type lectin-2, or dectin-2 (see AAW63010), that is essential for  
 CC DC-mediated T cell activation. Dectin-1 cDNA (see AAV42548) has  
 CC also been identified. The invention provides: dectin-1 and

CC dectin-2 polypeptides (see also AAW63009-22 and AAW69236-37), useful  
 CC for purifying T cells, for detecting autoantibodies and for  
 CC up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see  
 CC AAV42548-54, AAV42558-60 and AAV44850-51); expression vectors;  
 CC recombinant host cells; probes and primers useful e.g. for identifying  
 CC human dectin-like molecules (see AAV42560); antibodies; compounds that  
 CC modulate dectin-mediated activation of T cells; transgenic animals  
 CC useful for studying dectin function and for drug screening; and  
 CC dectin ligands. Dectin expression can be downregulated by  
 CC antisense sequences or ribozymes, or by inactivating the dectin  
 CC genes by homologous recombination, e.g. abrogation of dectin  
 CC expression is useful for treating allergy and autoimmune  
 CC disease. Alternatively gene disruption is done in vitro, the  
 CC cells pulsed with an antigen, and returned to the patient to  
 CC provide long-term non-responsiveness to the antigen. Since DC are  
 CC 'professional' antigen-presenting cells, genetic vaccines targeted  
 CC to them should not induce tolerance.  
 XX  
 SQ Sequence 1227 BP; 335 A; 271 C; 276 G; 345 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,82e-76 Length: 1227  
 Score: 793.00 Matches: 145  
 Percent Similarity: 77.00% Conservative: 19  
 Best Local Similarity: 68.08% Mismatches: 41  
 Query Match: 68.07% Indels: 8  
 DB: 19 Gaps: 4  
 US-09-766-511b-53 (1-209) x AAV42549 (1-1227)  
 QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
 Db 146 ATGGTGCAGGAAGACAAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193  
 QY 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuSerAlaCysPheIleValSer 38  
 Db 194 AGACTCTGGTCAGCTGCTGTGATTTCCATGTTACTCTTGAGTACCTGTTTCATTCGGAGC 253  
 QY 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
 Db 254 TGTGTGGTGACTTACCAATTTATTATGGACCCAGTGAAGACTATATGAACCTCAC 313  
 QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
 Db 314 ACATACCATTTCCAGTCTCACCTCTTCAGTGAAGGAGCTATGCTGCAGAAAAATGTGG 373  
 QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGlu 96  
 Db 374 GGATGCTGCCAAATCACTGGAAGTCACTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 433  
 QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
 Db 434 GAGAACTTCTGGAGCACCAGTGCAGAGAACCTGTTCAGATGGGGCTCATCTGGTGGTG 493  
 QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
 Db 494 ATCAATACTGAAGCGGAGCAGAAATTCATACCAGCAGCTGAATGAGTCACTTCTTAC 553  
 QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 Db 554 TTCTTGGTCTTTCGGATCCACAGGTAAATGGCAATGGCAATGGATGATGATCTCCT 613  
 QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 Db 614 TTCAGTCAAAATGTCAGGTTCCTGGCACCCCATGACCCCAATCTCCAGAGAGCGGTGT 673  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 Db 674 GTTTCATAGTTTACTGGAATCCTTCGAAATGGGGCTGGAATGATGTTTCTGTGATAGT 733  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 734 AAACACAATTCATATGTGAAATGAAGAAGATTACCTA 772



XX 15-MAR-2001.  
XX 30-JUN-2000: 2000WO-US18174.  
XX 10-SEP-1999: 99US-0393996.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI: 2001-183280/18.  
XX P-PSDB; AAU00480.  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX  
XX Disclosure; Fig 6E-6F; 326pp; English.  
XX  
CC The present sequence encoding for murine TANGO 405 protein is isolated  
CC from cDNA clone jtmMa025a11 from a long-term bone marrow cDNA library.  
CC An alternative splice variant of this protein is also described  
CC (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
CC INTERCEPT 400 (AAU00477). Six novel human proteins which include  
CC TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
CC INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
CC (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
CC described. The nucleic acids encoding these novel proteins are useful as  
CC modulating agents in regulating a variety of cellular processes and can  
CC be used to express the proteins in a host cell in gene therapy  
CC applications. Human and murine TANGO 405 proteins show sequence homology  
CC to murine lectin-2. TANGO 405 modulates growth, proliferation, survival,  
CC differentiation, activity, morphology and movement/migration of human  
CC lymphocytes and bone marrow cells and tissues and can be used to prevent,  
CC diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
XX  
XX Sequence 821 BP; 214 A; 188 C; 195 G; 223 T; 1 other;  
SQ

Alignment Scores:  
Pred. No.: 9,09e-71 Length: 821  
Score: 740.00 Matches: 142  
Percent Similarity: 75.94% Conservative: 19  
Best Local Similarity: 66.98% Mismatches: 42  
Query Match: 63.52% Indels: 10  
DB: 22 Gaps: 4

US-09-766-511B-53 (1-209) x AAS01376 (1-821)

QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
Db 174 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCGTG 221  
QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
Db 222 AGACTGTGTGACGTCTGTGATTTCCATGTTACTCTTGAGTACCTGTTTCATTGCGGAG 281  
QY 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
Db 282 TGTGTGGTGACCTACCAATTATTATGACCCAGCCAGCCATAGAGACTATATGACTTCAC 341  
QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
Db 342 ACATACCAATTCAGTCTCACTGCTTCACTGAGTGAAGGACTATGGTGTGAGAAAAATGTGG 401  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96  
Db 402 GGATGCTGCCAAATCACTGGAGTCAITTTGGCTCCAGCTACTCCTCATTTTACCAAG 461  
QY 97 GluLysValTrpSerLysSerGluGlnAsnGlnValGluMetGlyAlaHisValVal 116  
Db 462 GAGAACTTCTGGAGCACCAGTCAGCAGAACTCTGTTCAGATGGGGCTCATCTGTGGTGTG 521

QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGlnSerPheSerTyr 136  
Db 522 ATCAATACTGAAGCGGAGCAGAAATTTCAATCCAGCAGCTGAATGAGTCACTTCTTAC 581  
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
Db 582 TTCCTGGGTCTTCGGATCCCAA-GGTAATGGCAATGGCAATGGATCGATGATCTCCT 640  
QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
Db 641 TTCAGTCAAAATGTCAGTTCTGGCACCCCATCAACCCCAATCTTCAGAGAGCGGTGT 700  
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrp-AsnAspValIleCysGluTh 196  
Db 701 GTTTCAATAGTTTACTGGAATCCTTCGAAATGGGCTGGGAATCATCTTTCTGTGATAG 760  
QY 196 rArgArgAsnSerIleCysGluMetAsnLysIle 207  
Db 761 TAAACACAATTCATATGTGAAATGAANAAGATT 794

RESULT 7  
AAV42554  
ID AAV42554 standard; cDNA; 501 BP.  
XX  
AC AAV42554;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Mouse lectin-2 extracellular domain cDNA.  
XX  
KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
KW drug screening; ss.  
XX  
OS Mus sp.  
XX  
PN W09828332-A2.  
XX  
PD 02-JUL-1998.  
XX  
PF 22-DEC-1997; 97WO-0523761.  
XX  
PR 20-DEC-1996; 96US-0772440.  
XX  
PA (TEXA ) UNIV TEXAS SYSTEM.  
XX  
PI Arizumi K, Takashima A;  
XX  
DR WPI: 1998-377594/32.  
DR P-PSDB; AAW63022.  
XX  
PT Nucleic acid encoding dendritic cell specific peptide(s) lectin-1  
PT and -2 - useful, e.g. to regulate immune response, as vaccine  
PT adjuvants, for diagnosis and drug screening  
XX  
PS Disclosure; Page 153; 200pp; English.  
XX  
CC This cDNA sequence codes for the extracellular domain (see  
CC AAW63022) of a novel murine dendritic cell (DC) specific member of  
CC the c-type lectin family, termed DC-associated c-type lectin-2, or  
CC lectin-2 (see also AAW63010), that is essential for DC-mediated T  
CC cell activation. Full-length cDNA (see AAV42549) was isolated by  
CC subtractive hybridisation from BALB/c mouse epidermis DC line XS52.  
CC The invention provides: lectin-1 and -2 polypeptides (see also  
CC AAW63009-22 and AAW69236-37), useful for purifying T cells, for  
CC detecting autoantibodies and for up-regulating immunity e.g. as  
CC vaccine adjuvants; lectin DNA (see AAV42548-54, AAV42558-60 and  
CC AAV44850-51); expression vectors; recombinant host cells; probes  
CC and primers; antibodies; compounds that modulate lectin-mediated  
CC activation of T cells; transgenic animals; and lectin ligands.  
CC Dectin expression can also be down-regulated to treat allergy and  
CC autoimmune disease. Recombinant extracellular domain has been  
CC expressed as a His-tagged protein (see AAW69237) useful for antibody

CC production.  
XX  
SQ Sequence 501 BP; 142 A; 108 C; 115 G; 136 T; 0 other;

Alignment Scores:  
Pred. No.: 1.9e-64 Length: 501  
Score: 679.00 Matches: 118  
Percent Similarity: 79.04% Conservatve: 14  
Best Local Similarity: 70.66% Mismatches: 33  
Query Match: 58.28% Indels: 2  
DB: 19 Gaps: 1

US-09-766-511b-53 (1-209) x AAV42554 (1-501)

Qy 45 PheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeu 64  
Db 1 TTTATTATGGACCGCCAGTGAAGACTATATGACTTCACACATACCATCCAGTC1C 60  
Qy 65 ThrCysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSer 82  
Db 61 ACCTGCTTCAGTGAAGGGACTATGGTGTGCAGAAAAAATGGGGGATGCTGCCCAAAATCAC 120  
Qy 83 TrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLys 102  
Db 121 TGGAAAGTCATTGGCTCCAGCTGCTACCTCATTTCTACCAAGGAGAACTTCTGGAGCAC 180  
Qy 103 SerGluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGlu 122  
Db 181 AGTGAGCAGAACTGTGTTCAGATGGGGGCTCATCTGGTGTGATCAATACTGAAGCGGAG 240  
Qy 123 GlnAsnPheLeValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAsp 142  
Db 241 CAGAAATTCATCACCAGCAGCTGAATGAGTCACATTCTTACTTCCTGGGTCTTTTCGGAT 300  
Qy 143 ProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArg 162  
Db 301 CCACAAGGTAATGGCAATGGCAATGGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Qy 163 PheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrp 182  
Db 361 TTCTGGCACCCCATGAACCAATCTCCAGAGAGCGGTGTGTTCAATAGTTTACTGG 420  
Qy 183 LysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202  
Db 421 AATCCTTCGAAATGGGCTGGAATGATGTTTCTGTGATAGTAAACACAAATTCATATGT 480  
Qy 203 GluMetAsnLysIleTyrLeu 209  
Db 481 GAAATGAAGAAGATTTACCTA 501

RESULT 8  
AAF90241  
ID AAF90241 standard; cDNA; 1312 BP.

XX AAF90241;  
AC AAF90241;  
XX  
XX  
XX 06-AUG-2001 (first entry)

DE Nucleotide sequence of BDCA-2 antigen.

XX BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;  
KW viral infection; autoimmune disease; allergic response; cancer; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 491..1132  
FT CDS /\*tag= a  
FT /\*product= "BDCA-2"

XX WO2001136487-A2.

XX  
XX  
PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-IB01832.

XX 15-NOV-1999; 99US-0165555.

PR 23-NOV-1999; 99US-0167076.

PR 28-JAN-2000; 2000US-0179003.

PR 07-FEB-2000; 2000US-0180775.

PR 11-APR-2000; 2000US-0196824.

PR 13-APR-2000; 2000US-0197205.

XX (MILT-) MILTENYI BIOTECH GMBH.

XX Schmitz J, Dzionek A, Buck DW;

XX WPI; 2001-355622/37.

DR P-PSDB; ABB84215.

XX Compositions and cell populations enriched in dendritic cells through use of antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4, are used to treat viral infections, autoimmune diseases, allergic responses and cancer -

XX Claim 134; Fig 12; 115pp; English.

XX The present sequence encodes an antigen designated BDCA-2. The specification describes compositions and hematopoietic cell populations enriched in dendritic cells (DCs). These compositions are produced using antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4. The DCs obtained are used to treat viral infections, autoimmune disease, allergic response, and cancer. BDCA-1, BDCA-2, BDCA-3 and BDCA-4 monoclonal antibodies and their antigen-binding fragments are used to detect, enumerate and isolate DC populations from leukapheresis material, whole blood and tonsils and from non-hematopoietic and hematopoietic tissues.

XX Sequence 1312 BP; 374 A; 301 C; 297 G; 340 T; 0 other;

Alignment Scores:  
Pred. No.: 2.81e-59 Length: 1312  
Score: 637.00 Matches: 115  
Percent Similarity: 68.08% Conservatve: 30  
Best Local Similarity: 53.99% Mismatches: 64  
Query Match: 54.68% Indels: 4  
DB: 22 Gaps: 2

US-09-766-511b-53 (1-209) x AAF90241 (1-1312)

Qy 1 MetMetGlnGluGlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArg 19  
Db 491 ATGGTGCCTGAAGAAGAGCGCTCAAGACCGAGAGAGGACTCTGGTGGTTCCAGTTGAAG 550

Qy 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39  
Db 551 GTCGTGTCCTGGCAGCTCGTATCCATCTGCTCCTCAGTCTCTGTTCTACTGTAGTCT 610

Qy 40 ValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57  
Db 611 GTGGTGCCTCACAAATTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAGTTACGAGAG 670

Qy 58 ---HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76  
Db 671 TATCAACAGCATATCATCCCAAGCCTGACCTGCGTCATGGAAGAGAACACATAGAAGATTGG 730

Qy 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96  
Db 731 AGCTGTCGCCCAACCCCTTGGACTTCATTTTCAGTCTAGTGTCTACTTTATTCTACTGGG 790

Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
Db 791 ATGCAATCTGGCACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGGTG 850

Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
Db 851 ATCAACACCCAGGGAAGAACAGGATTTTCATCATTCAGAATCTCAAAAGAAATCTTCTTAT 910

QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 DB 911 TTTTGGGGCTGTCAGATCCAGGGGTCGGCAGATGGCAATGGTTGACCAGACCA 970  
 QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 DB 971 TACAATGAAATGTACATCTCGGCACCTCAGGTGAACCAATAACCTTGATGAGCGTTGT 1030  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 DB 1031 GCGAATAATAATTCGTTCTTCAGAGAATGGGGCTGGAATGCATTCACCTGTCATGTA 1090  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 DB 1091 CCTCAGAGTCAATTTGCAAGATGAAGAAGATCTACATA 1129  
 RESULT 9  
 AAD19729  
 ID AAD19729 standard; DNA; 827 BP.  
 XX  
 AC AAD19729;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Dendritic cell (DC) DCLEC gene.  
 XX  
 KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
 KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
 KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
 KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
 KW transplant rejection; chronic inflammatory disease; anti-HIV; ds.  
 XX  
 OS Unidentified.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT CDS  
 FT 1..657  
 FT /\*tag= a  
 FT /product= "Dendritic cell DCLEC protein"  
 XX  
 PN WO200172773-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PF 28-MAR-2001; 2001WO-EP03542.  
 XX  
 PR 29-MAR-2000; 2000US-192934P.  
 PR 18-MAY-2000; 2000US-205020P.  
 PR 18-MAY-2000; 2000US-205026P.  
 PR 19-MAY-2000; 2000US-205767P.  
 PR 19-MAY-2000; 2000US-205769P.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 XX  
 PI Werner G, Phares W, Jaritz M, Lapp H, Kalthoff rs;  
 XX  
 DR WPI; 2001-616466/71.  
 DR P-PSDB; AAEL2079.  
 XX  
 PR New polypeptides for screening therapeutic agonists and antagonists  
 PT comprise dendritic cell polypeptides -  
 XX  
 PS Claim 1; Page 50-51; 52pp; English.  
 XX  
 CC The invention relates to dendritic cell (DC) proteins and their  
 CC corresponding DNA molecules. A pharmaceutical composition comprising  
 CC agonist and antagonist of DC proteins are useful for treating abnormal  
 CC conditions related to both an excess of and insufficient level of  
 CC expression of DC gene, or related to both an excess of and insufficient  
 CC activity of DC protein. Soluble form of DC proteins are used as an active  
 CC ingredient in combination with pharmaceutical acceptable carriers.  
 CC DC genes and proteins are useful for treating chronic inflammatory

CC diseases, autoimmune diseases, transplant rejection crisis, including  
 CC inflammatory skin diseases such as contact hypersensitivity, atopic  
 CC dermatitis or virally-induced immune suppression such as AIDS and cancer.  
 CC DC protein is useful for inducing immunological response in a mammal, and  
 CC as immunogen to produce antibodies immunospecific for the polypeptide.  
 CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
 CC reagent, and for chromosomal identification. The present sequence is  
 CC dendritic cell (DC) DCLEC gene which is found to belong to the family  
 CC of C-type lectins with one single carbohydrate recognition domain at the  
 CC C-terminal end.  
 XX  
 SQ Sequence 827 BP; 242 A; 164 C; 188 G; 233 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8 31e-59 Length: 827  
 Score: 630.00 Matches: 114  
 Percent Similarity: 67.94% Conservatives: 28  
 Best Local Similarity: 54.55% Mismatches: 63  
 Query Match: 54.08% Indels: 4  
 DB: 22 Gaps: 2  
 US-09-766-511B-53 (1-209) x AAD19729 (1-827)  
 QY 5 GlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArgLeuTrpSerVal 23  
 DB 28 CAAGAGCCTCAGACCGCAGAGAAAGGACTCTGGTGTTCAGTTGAAGCTCTGGTCCATG 87  
 QY 24 AlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyr 43  
 DB 88 GCAGTCGTATCCATCTGCTCCCTCAGTGTCTGTTTCTCAGTGTCTGTGGTCCCTCAC 147  
 QY 44 HisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSerTyr 60  
 DB 148 AATTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAAGTTACGACAGATCAACAGTAT 207  
 QY 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
 DB 208 CATCCAAGCCTGACCTGCGTCTCATGGAAGAAAGACATAGAAATTGGAGCTGCTGCCCA 267  
 QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
 DB 268 ACCCTTGGACTTCATTTCTAGTCTAGTCTTCTACTTTTCTACTGGGATGCAATCTTGG 327  
 QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
 DB 328 ACTAAGAGTCAAAAGAACTGTCTCTGTGATGGGGGCTGATCTGTGTGATCAACACAGG 387  
 QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
 DB 388 GAAGAACAGGATTTTCATCTCAGAAATCTGAAAGAAATTTCTTTATTTCTGGGGCTG 447  
 QY 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
 DB 448 TCAGATCCAGGGGTCGGCAGATTCGCAATGGTTTGACCAGACACCATACATATAAAT 507  
 QY 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal 180  
 DB 508 GTCACATTTCTGGCACCCTCAGGTGAACCAATAACCTTGTATGAGCGTTGTGCGATAAAT 567  
 QY 181 PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer 200  
 DB 568 TTCCTGTTCTCAGAGAATGGGGCTGGAATGCATTCACCTGTCTACCTGTCTCAGAGTCA 627  
 QY 201 IleCysGluMetAsnLysIleTyrLeu 209  
 DB 628 ATTTGCAAGATGAAGAAGATCTACATA 654  
 RESULT 10  
 AAD19730  
 ID AAD19730 standard; DNA; 800 BP.  
 XX  
 AC AAD19730;  
 XX



18-DEC-2001 (first entry)  
Dendritic cell (DC) DCLEC/SPLICE 1 variant gene.  
Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
transplant rejection; chronic inflammatory disease; anti-HIV; variant;  
ds.  
Unidentified.  
Key Location/Qualifiers  
CDS 1..634  
/\*tag= a  
/product= "Dendritic cell DCLEC/SPLICE 1 variant protein"  
/trans\_except= (pos:1, aa:Pro)  
/note= "This codon has an apparent deletion of 2  
nucleotides, which alters the reading frame; CDS does not  
include start codon"  
/partial  
W0200172773-A2.  
04-OCT-2001.  
28-MAR-2001; 2001WO-EP03542.  
29-MAR-2000; 2000US-192934P.  
18-MAY-2000; 2000US-205020P.  
18-MAY-2000; 2000US-205026P.  
19-MAY-2000; 2000US-205767P.  
19-MAY-2000; 2000US-205769P.  
(NOVS ) NOVARTIS AG.  
(NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
WPI: 2001-616466/71.  
P-PSDB; AAEI2080.  
New polypeptides for screening therapeutic agonists and antagonists  
comprise dendritic cell polypeptides -  
Claim 1; Page 51; 52pp; English.  
The invention relates to dendritic cell (DC) proteins and their  
corresponding DNA molecules. A pharmaceutical composition comprising  
agonist and antagonist of DC proteins are useful for treating abnormal  
conditions related to both an excess of and insufficient level of  
expression of DC gene, or related to both an excess of and insufficient  
activity of DC protein. Soluble form of DC proteins are used as an active  
ingredient in combination with pharmaceutical acceptable carriers.  
DC genes and proteins are useful for treating chronic inflammatory  
diseases, autoimmune diseases, transplant rejection crisis, including  
inflammatory skin diseases such as contact hypersensitivity, atopic  
dermatitis or virally-induced immune suppression such as AIDS and cancer.  
DC protein is useful for inducing immunological response in a mammal, and  
as immunogen to produce antibodies immunospecific for the polypeptide.  
DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
reagent, and for chromosomal identification. The present sequence is  
dendritic cell (DC) DCLEC/SPLICE 1 variant gene which is found to belong  
to the family of C-type lectins with one single carbohydrate recognition  
domain at the C-terminal end.  
Sequence 800 BP; 231 A; 160 C; 183 G; 226 T; 0 other;  
Alignment Scores:  
Pred. No.: 2,05e-57 Length: 800  
Score: 617.00 Matches: 111  
Percent Similarity: 67.14% Conservative: 30

Best Local Similarity: 52.86% Mismatches: 63  
Query Match: 52.96% Indels: 6  
DB: 22 Gaps: 2  
US-09-766-511b-53 (1-209) x AAD19730 (1-800)  
QY 6 GlnProGlnSerThrGluLysArg-----GlyTrpLeuSerLeuArgLeuTrpSer 22  
Db 2 CGCCCGCGCTCCGAGACCGAGAGAAAGGACTCTGGTGGTTCCAGTTGAAGCTCGGTCC 61  
QY 23 ValAlaGlyIleSerIleAlaLeuSerAlaCysPheIleValSerCysValValThr 42  
Db 62 ATGGCAGTCGTATCCATCTTGCCTCAGTCTGTCTTTCACGTGTGAGTTCTGTGGTGCCT 121  
QY 43 TyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSer 59  
Db 122 CACAATTTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAGTTACGAGAGTATCAACAG 181  
QY 60 TyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCys 79  
Db 182 TATCATCCAAGCCTGACCTCGCTCATGGAAGGAAGACATAGAGATTGGAGCTGCTGCC 241  
QY 80 ProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysVal 99  
Db 242 CCAACCCCTTGGACTTCATTTCAGTCTAGTTGCTTACTTATTTCTACTGGGATGCAATCT 301  
QY 100 TrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThr 119  
Db 302 TGGACTAAGAGTCAAAAGAACTGTTCTGTGTGGGGGCTGATCTGGTGGTGTATCAACACC 361  
QY 120 GluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGly 139  
Db 362 AGGGAAGAACAGGATTTTCATCATTCAGATCTGAAAGAAATCTCTTATTTCTGGGG 421  
QY 140 LeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLys 159  
Db 422 CTGTCAAGTCCAGGGGGTGGCGACATTGGCAATGGTGTGACACACACATACAATGAA 481  
QY 160 AsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIle 179  
Db 482 AATGTCACATCTTGGCAGCTCAGGTGAACCAATAACCTTGTATGATGCGGATAATA 541  
QY 180 ValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsn 199  
Db 542 AATTTCCGTTCTTCAGAGAATGGGGCTGGAATGATGATCATCTGTCATGCTCAGAAAG 601  
QY 200 SerIleCysGluMetAsnLysIleTyrLeu 209  
Db 602 TCAATTTGCAAGATGAAGAAGATCTACATA 631  
RESULT 11  
ABK52914  
ID ABK52914 standard; cDNA; 444 BP.  
XX  
AC ABK52914;  
XX  
DT 27-AUG-2002 (first entry)  
XX Human cDNA encoding a partial dendritic cell immunoreceptor #2.  
DE Human; ss; gene; dendritic cell immunoreceptor; cytostatic;  
KW antitachymatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;  
KW asthma; inflammation; obesity; diabetes; central nervous system disorder;  
KW Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;  
KW cardiovascular disorder; myocardial infarction; ischaemic heart disease;  
KW congestive heart failure; chronic obstructive pulmonary disease; COPD.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 1..444  
FT /\*tag= a  
FT /product= "Dendritic cell immunoreceptor"



CC of the polypeptide or nucleic acid. The reagent is useful for the  
CC preparation of a medicament for modulating the activity of human  
CC dendritic cell immunoreceptor in a disease such as cancer, allergy,  
CC anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous  
CC system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or  
CC dementia), or a cardiovascular disorder (e.g. myocardial infarction,  
CC ischaemic heart disease, congestive heart failure), chronic obstructive  
CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
CC disorders are listed in the specification). The polypeptide and nucleic  
CC acid are useful for identifying test compounds which act as agonists or  
CC antagonists, for raising specific antibodies, and as a bait protein in a  
CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
CC diagnostic assays for detecting diseases and abnormalities or  
CC susceptibility to disease and abnormalities related to the presence of  
CC mutations. The present sequence is a cDNA encoding a partial  
CC dendritic cell immunoreceptor.

XX Sequence 402 BP; 114 A; 80 C; 96 G; 112 T; 0 other;

Alignment Scores:                      1.13e-40                      Length:                      402  
Pred. No.:                      459.00                      Matches:                      77  
Score:                      70.90%                      Conservative:                      18  
Percent Similarity:                      57.46%                      Mismatches:                      39  
Best Local Similarity:                      39.40%                      Indels:                      0  
Query Match:                      24                      Gaps:                      0  
DB:

US-09-766-511b-53 (1-209) x ABK52903 (1-402)

Qy 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95  
Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTCCTACTTTTCTACT 60  
Qy 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115  
Db 61 GGGATGCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG 120  
Qy 116 ValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSer 135  
Db 121 GTGATCAACACCGGGAGAGACAGGATTTTCATTCATGAACTCGAAGAAATCTTCTCI 180  
Qy 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155  
Db 181 TATTTCTGGGCTGTGCAGATCCAGGGGTCGGCGACATTTGGCAATGGTGTGACCACACA 240  
Qy 156 ProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln 175  
Db 241 CCATACATGAAATGTGCATTTCTGGCACTTCAGGTGAACCCCAATAACCTTGATGAGCGT 300  
Qy 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195  
Db 301 TGTGGGATAATAATTTCCGTTCTTCAGAGAATGGGGCTGGGAATGACATTCATGTCAT 360  
Qy 196 ThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
Db 361 GTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 402

RESULT 13  
AAS31385  
ID AAS31385 standard; cDNA; 1091 BP.  
XX  
AC AAS31385;  
XX  
DF 04-DEC-2001 (first entry)  
DE Human cDNA encoding a novel extracellular matrix protein, Seq ID No 199.  
XX  
KW Human: secreted extracellular matrix protein; ss: immunomodulatory;  
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antiaizheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;

KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX Homo sapiens.  
XX WO200155368-A1.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01348.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 16-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236602.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI P-PSDB; RAU19814.  
 XX WPI; 2001-465572/50.  
 DR P-PSDB; RAU19814.  
 XX Nucleic acid molecules encoding human secreted extracellular matrix  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers -  
 XX Claim 1; SEQ ID No 199; 577pp; English.  
 PS The invention relates to isolated nucleic acid molecules encoding  
 CC novel human secreted extracellular matrix proteins (SPs). The  
 CC polynucleotides and proteins are used to prevent, treat a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. For example, disorders associated with decreased  
 CC expression of SPs. The SP polynucleotide or a vector expressing them may  
 CC be administered to treat diseases by gene therapy. Antisense molecules  
 CC may be administered to down regulate expression of SPs by binding with  
 CC the cells own genes and preventing their expression. The polynucleotides  
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be  
 CC used as antigens to produce antibodies and to identify modulators  
 CC (agonists and antagonists) of the SPs. The anti-SP antibodies and  
 CC antagonists may also be used to down regulate expression and activity of  
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.  
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac  
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses  
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses  
 CC include wound healing, maintenance of organs before transplantation,  
 Alignment Scores:  
 Pred. No.: 6.16e-40 Length: 1091  
 Score: 458.00 Matches: 83  
 Percent Similarity: 62.57% Conservative: 34  
 Best Local Similarity: 44.39% Mismatches: 66  
 Query Match: 39.31% Indels: 4  
 DB: 22 Gaps: 3  
 US-09-766-511b-53 (1-209) x AAS31385 (1-1091)  
 QY 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45  
 Db 172 ATATTTTCTGCTATTTGGCAATCTCATTTCTTTTGTTCATTTCTTTTCAA--- 228  
 QY 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65  
 Db 229 AAATATTTCTCAGCTTCTTGAAAAAAGACATACAAAAGAGCTGGTTCATACAACTGGAG 288  
 QY 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83  
 Db 289 TGTGTGAAAAAATAATGCCCCGTGGAGAGACAGCCCTGGAGCTGTTGCCCAAGAAATTGG 348  
 QY 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103  
 Db 349 AAGTCATTTAGTTCCTCAACTGCTACTTTATTTCTTACCTGACATCTTGGCAAGACAGT 408  
 QY 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123  
 Db 409 GAGAAAGGACTGTGCTAGATGGAGGCTCACCTGCTGTGATATAACACTCAAGAAGAGACAG 468  
 QY 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143  
 Db 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143

Db 469 GATTTCATCTTCAGAAATCTGCAAGAAGAAATCTGCTATTATTTGTGGGGCTCTCAGATCCA 528  
QY 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
Db 529 GAAGGTCAAGGACATTGGCAATGGTGTGATCAGACACCAATGAAAGTCCACATTC 588  
QY 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183  
Db 589 TGGCATCCACGTGAGCCAGTGCATCCCAATGAGCGTGGCTTGTGCTAAATTTTCGTAA 648  
QY 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202  
Db 649 TCACCAAAAGATGGGGCTGGAATGATTAATGTGCTTGGTCCCAAGGTCAGTTTGT 708  
QY 203 GluMetAsnLysIleTyrLeu 209  
Db 709 GAGATGATGAAGATCCACTTA 729  
RESULT 14  
ABQ66709  
ID ABQ66709 standard; cDNA; 1091 BP.  
XX AC ABQ66709;  
XX XX  
XX 23-AUG-2002 (first entry)  
XX XX  
XX Human polynucleotide SEQ ID NO 199.  
XX XX  
XX Human; nootropic; neuroprotective; cytostatic; dermatologic; virocidic;  
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; nephrotropic; gene therapy; vaccine;  
XX gene; ss.  
XX  
XX Homo sapiens.  
XX XX  
XX US2002042386-A1.  
XX XX  
XX 11-APR-2002.  
XX XX  
XX 17-JAN-2001; 2001US-0764870.  
XX XX  
XX 31-JAN-2000; 2000US-179065P.  
XX 04-FEB-2000; 2000US-180628P.  
XX 28-JUN-2000; 2000US-214886P.  
XX 07-JUL-2000; 2000US-216647P.  
XX 07-JUL-2000; 2000US-216880P.  
XX 11-JUL-2000; 2000US-217487P.  
XX 11-JUL-2000; 2000US-217496P.  
XX 14-JUL-2000; 2000US-218290P.  
XX 26-JUL-2000; 2000US-220963P.  
XX 26-JUL-2000; 2000US-220964P.  
XX 14-AUG-2000; 2000US-224518P.  
XX 14-AUG-2000; 2000US-224519P.  
XX 14-AUG-2000; 2000US-225267P.  
XX 14-AUG-2000; 2000US-225268P.  
XX 14-AUG-2000; 2000US-225270P.  
XX 14-AUG-2000; 2000US-225447P.  
XX 14-AUG-2000; 2000US-225757P.  
XX 14-AUG-2000; 2000US-225758P.  
XX 22-AUG-2000; 2000US-226868P.  
XX 30-AUG-2000; 2000US-228924P.  
XX 01-SEP-2000; 2000US-228928P.  
XX 01-SEP-2000; 2000US-229343P.  
XX 01-SEP-2000; 2000US-229344P.  
XX 01-SEP-2000; 2000US-229345P.  
XX 05-SEP-2000; 2000US-229509P.  
XX 05-SEP-2000; 2000US-229513P.  
XX 08-SEP-2000; 2000US-231413P.  
XX 21-SEP-2000; 2000US-234223P.

PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234997P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 29-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-239335P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241783P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 17-NOV-2000; 2000US-249299P.  
PR 08-DEC-2000; 2000US-251856P.  
PR 08-DEC-2000; 2000US-251868P.  
PR 08-DEC-2000; 2000US-251869P.  
XX  
XX (ROSE/) ROSEN C A.  
XX (RUBE/) RUBEN S M.  
XX (BARA/) BARASH S C.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
PI  
XX  
XX WPI; 2002-470713/50.  
XX P-PSDB; ABP48034.  
XX  
XX New nucleic acid encoding human proteins, useful for diagnosis,  
XX treatment and prevention of e.g. osteoporosis, also related  
XX polypeptides and antibodies  
XX  
XX Claim 1; SEQ ID NO 199; 235pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
XX (ABP47846-ABP48110) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from USPTO at seqdata.uspto.gov/sequence.html?docid=99909764870.  
XX  
XX Sequence 1091 BP; 308 A; 213 C; 250 G; 320 T; 0 other;

Alignment Scores:  
Pred. No.: 6.16e-40 Length: 1091  
Score: 458.00 Matches: 83  
Percent Similarity: 62.57% Conservative: 34  
Best Local Similarity: 44.39% Mismatches: 66  
Query Match: 39.31% Indels: 4  
DB: 24 Gaps: 3

US-09-766-511B-53 (1-209) x ABQ66709 (1-1091)

QY 26 ILeSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrThrHisPhe 45  
Db 172 ATATTTTCTCTGATATGGCAATCTCATCTTTATTCCTTTCTTTCTTTCTTTCAA--- 228

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65  
Db 229 AAATATCTCAGCTCTCTGAAAAAGACTACAAAGAGCTGGTTTCATACAACTGGAG 288  
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83  
Db 289 TGTGTGAAAAAATATGCGCGTGAAGACAGACCGCTGGAGCTGTGGCCCAAGAATGG 348  
Qy 84 LysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103  
Db 349 AAGTCATTAGTCTCACTGCTACTTTATTCTACTCAATCAGCATCTTGGCAAGACAGT 408  
Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123  
Db 409 GAGAAGGACTGTGTAGAAATGGAGGCTCAGCTGCTGTGATAAACACTCAAGAAGACAG 468  
Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143  
Db 469 GATTTCATCTCCAGATCTGCAGAGAAGATCTGCTTATTTGTGGGCTCTCAGATCCA 528  
Qy 144 GlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
Db 529 GAAGGTCAGGCACATTGGCAATGGTTGATCAGACACCATACATGAAGTTCACATTC 588  
Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183  
Db 589 TGGCATCCAGTGAGCCCATGATCCCAATGAGCGTGGCTGTGCTAAATTTTCGTAAA 648  
Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgAsnSerIleCys 202  
Db 649 TCACCCAAAGATGGGCTGGAATGATGTTAATTGCTTGGTCTCTCAAGGTCAGTTGT 708  
Qy 203 GluMetAsnLysIleTyrLeu 209  
Db 709 GAGATGATGAAGATCCACTTA 729  
RESULT 15  
AAS31224  
ID AAS31224 standard; cDNA; 1096 bp.  
XX AAS31224;  
AC AAS31224;  
DT 04-DEC-2001 (first entry)  
XX Human cDNA encoding a novel extracellular matrix protein, Seq ID No 38.  
XX Human; secreted extracellular matrix protein; ss; immunomodulatory;  
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX Homo sapiens.  
OS  
XX  
PN WO200155368-A1.  
XX  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01348.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.







GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 20, 2003, 18:37:42 : Search time 2230 seconds  
(without alignments)  
2727.572 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTKRRGWSLURL.....NDVICETRRNSICENMKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+p2n.model -DEV=vlh  
-Q/cgcn2\_1/USPTO.spool/US09766511/runat\_19022003\_094338\_29286/app.query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEL=0 -DOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09766511.ecgn.1.1.1616 @runat\_19022003\_094338\_29286 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	793	68.1	1227	6	AX155225	AX155225 Sequence
2	793	68.1	1227	10	AF240357	AF240357 Mus muscu
3	793	68.1	1254	10	BC023008	BC023008 Mus muscu
4	659	56.6	1125	10	AF240358	AF240358 Mus muscu
5	637	54.7	851	9	AF325459	AF325459 Homo sapi
6	637	54.7	1312	6	AX155223	AX155223 Sequence
7	637	54.7	1313	9	AF293615	AF293615 Homo sapi
8	630	54.1	827	6	AX357481	AX357481 Sequence
9	617	53.0	800	6	AX357483	AX357483 Sequence
10	592.5	50.9	1125	10	AF240359	AF240359 Mus muscu
11	534.5	45.9	739	9	AF325460	AF325460 Homo sapi
12	459.5	39.4	444	6	AX456975	AX456975 Sequence
13	459	39.4	402	6	AX456962	AX456962 Sequence
14	458	39.3	1133	9	AF109146	AF109146 Homo sapi
15	458	39.3	1271	9	HSAL133532	AJ133532 Homo sapi
16	453	38.9	1305	9	AF067800	AF067800 Homo sapi
17	445.5	38.2	1195	6	AX400064	AX400064 Sequence
18	445	38.2	1308	9	AF328684	AF328684 Homo sapi
19	445	38.2	1973	9	BC032313	BC032313 Homo sapi
20	443	38.0	918	10	AF061272	AF061272 Mus muscu
21	441	37.9	753	9	AF411850	AF411850 Homo sapi
C	439.5	37.7	582	6	AX456965	AX456965 Sequence
	436	37.4	1037	10	AF387099	AF387099 Mus muscu
23	436	37.4	1418	10	MMU133533	AJ133533 Mus muscu
24	436	37.4	615	9	AF200738	AF200738 Homo sapi
25	433.5	37.2	938	10	BC006623	BC006623 Mus muscu
26	430	36.9	568	6	AX456969	AX456969 Sequence
C	427.5	36.7	568	6	AX456969	AX456969 Sequence
	419.5	36.0	1192	10	BC034893	BC034893 Mus muscu
29	419	36.0	558	6	AX456968	AX456968 Sequence
C	400.5	34.4	403	6	AX456970	AX456970 Sequence
	398.5	34.2	936	9	BC000715	BC000715 Homo sapi
31	398.5	34.2	968	6	AX083498	AX083498 Sequence
32	398.5	34.2	997	6	AX375956	AX375956 Sequence
33	398.5	34.2	2143	9	AB024718	AB024718 Homo sapi
34	398.5	34.2	1606	10	BC003218	BC003218 Mus muscu
35	390.5	33.5	2517	10	AB024717	AB024717 Mus muscu
36	382	32.8	1013	6	AX456966	AX456966 Sequence
C	377.5	32.4	753	6	AX456967	AX456967 Sequence
	375.5	32.2	561	6	AX456972	AX456972 Sequence
40	299.5	25.7	448	6	AX456973	AX456973 Sequence
41	295.5	25.4	117296	9	AC092865	AC092865 Homo sapi
C	290.5	24.9	666	6	AX456971	AX456971 Sequence
	289.5	24.8	109320	9	AC092746	AC092746 Homo sapi
C	267.5	23.0	559	6	AX456974	AX456974 Sequence
	259	22.2	1246	10	MMHL2	X53042 Mouse MHL-2

ALIGNMENTS

RESULT 1

AX155225  
LOCUS AX155225 1227 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 3 from Patent WO0136487.  
ACCESSION AX155225  
VERSION AX155225.1 GI:14536703  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Schmitz, J., Dzionek, A. and Buck, D.W.  
TITLE Antigen-binding fragments specific for dendritic cells, compositions and methods of use thereof antigens recognized thereby and cells obtained thereby  
JOURNAL Patent: WO 0136487-A 3 25-MAY-2001;  
FEATURES Miltenyi Biotec GmbH (DE)  
source Location/Qualifiers  
1..1227  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
146..775  
/note="coding sequence of mouse Dectin-2"  
/codon\_start=1  
/protein\_id="CAC42676.1"  
/db\_xref="GI:14536704"  
/translation="MVQEROSQGVWTLRLWSAAVISMILLSTCFIASCVVYQFI  
MDQSRRLYELHTVHSSLCFSEGTVMSEKMWGCPNHWKSGSCYLSTKKNFWS  
SEQNCVOMGAHLVINTAEQNFITQOLNESLSYFLGLSDPQSGKQWIDTFFSQN  
VRFWHPHPNLPERCVSIVYWNPSKGNWDFCDSKHNSICEMKKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,97e-73 Length: 1227  
Score: 793.00 Matches: 145  
Percent Similarity: 77.00% Conservative: 19  
Best local Similarity: 68.08% Mismatches: 41  
Query Match: 68.07% Indels: 8  
DB: 6 Gaps: 4  
US-09-766-511B-53 (1-209) x AX155225 (1-1227)  
QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
Db 146 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193  
QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
Db 194 AGACTCTGTCAGCTGCTGTGATTTCCATGTTACTCTTTGAGTACTCTGTTTCATTGCGAGC 253  
QY 39 CysValValThrTyrrHisPheThrTyrrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
Db 254 TGTGTGGTGACTTACCAATTTATGATGACAGCCAGCCAGTAGAAGACTATATGAACCTCAC 313  
QY 59 SerTyrrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
Db 314 ACATACCAATTCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGTCAAGAAAAATGTGG 373  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrrPheIleSerSerGlu 96  
Db 374 GGATGCTGCCAAATACACTGGAAGTCAITTTGGCTCCAGCTGCTACCTCATTTTCACCAAG 433  
QY 97 GlyLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
Db 434 GAGAACTTCTGGAGCACCAGTGACAGAACTGTGTTTCAGATGGGGGCTCATCTGCTGGTG 493  
QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuLeuAsnGluSerPheSerTyrr 136  
Db 494 ATCACTACTGAAGCGAGCAGCAATTTTCATCACCCAGCAGCTGAATGATGACTTCTTAC 553  
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
Db 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156

Db 554 TTCTCTGGGTCTTTTCGGATCCACAAAGGTAATGGCAATGGCAATGGATCGATGATCTCT 613  
QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
Db 614 TTCAGTCAAAATGTCAGGTCTTGGACCCCATGAACCAATCTTCCAGAGAGCGGTGT 673  
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
Db 674 GTTTCATAGTTTACTTGGGAATCTTCGAAATGGGCTGGAATGATGTTTCTGTGATAGT 733  
QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrrLeu 209  
Db 734 AAACACAATTCATATGTGAATGAAGAAGATTACCTA 772  
RESULT 2  
AF240357  
LOCUS AF240357 1227 bp mRNA linear ROD 02-MAY-2000  
DEFINITION Mus musculus strain BALB/c dectin-2 alpha isoform mRNA, complete cds.  
ACCESSION AF240357  
VERSION AF240357.1 GI:7677468  
KEYWORDS Mus musculus.  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Ariizumi, K., Shen, G.-L., Shikano, S., Ritter, R. III, Zukas, P., Edelbaum, D., Morita, A. and Takashima, A.  
TITLE Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms  
JOURNAL J. Biol. Chem. (2000) In press  
REFERENCE 2 (bases 1 to 1227)  
AUTHORS Ariizumi, K. and Takashima, A.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA  
FEATURES Location/Qualifiers  
1..1227  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="XS2"  
/cell\_type="dendritic cell"  
146..775  
/note="C-type lectin; type II transmembrane protein; dendritic cell-associated expression; C-type lectin motif"  
/codon\_start=1  
/product="dectin-2 alpha isoform"  
/protein\_id="AAF67177.1"  
/db\_xref="GI:7677469"  
/translation="MVQEROSQGVWTLRLWSAAVISMILLSTCFIASCVVYQFI  
MDQSRRLYELHTVHSSLCFSEGTVMSEKMWGCPNHWKSGSCYLSTKKNFWS  
SEQNCVOMGAHLVINTAEQNFITQOLNESLSYFLGLSDPQSGKQWIDTFFSQN  
VRFWHPHPNLPERCVSIVYWNPSKGNWDFCDSKHNSICEMKKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,97e-73 Length: 1227  
Score: 793.00 Matches: 145  
Percent Similarity: 77.00% Conservative: 19  
Best Local Similarity: 68.08% Mismatches: 41  
Query Match: 68.07% Indels: 8  
DB: 10 Gaps: 4  
US-09-766-511B-53 (1-209) x AF240357 (1-1227)  
QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
Db 146 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193  
QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
Db 194 AGACTCTGTCAGCTGCTGTGATTTCCATGTTACTCTTTGAGTACTCTGTTTCATTGCGAGC 253  
QY 39 CysValValThrTyrrHisPheThrTyrrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
Db 254 TGTGTGGTGACTTACCAATTTATGATGACAGCCAGCCAGTAGAAGACTATATGAACCTCAC 313  
QY 59 SerTyrrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
Db 314 ACATACCAATTCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGTCAAGAAAAATGTGG 373  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrrPheIleSerSerGlu 96  
Db 374 GGATGCTGCCAAATACACTGGAAGTCAITTTGGCTCCAGCTGCTACCTCATTTTCACCAAG 433  
QY 97 GlyLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
Db 434 GAGAACTTCTGGAGCACCAGTGACAGAACTGTGTTTCAGATGGGGGCTCATCTGCTGGTG 493  
QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuLeuAsnGluSerPheSerTyrr 136  
Db 494 ATCACTACTGAAGCGAGCAGCAATTTTCATCACCCAGCAGCTGAATGATGACTTCTTAC 553  
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
Db 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156

```

Db 194 AGACTCTGGTCAGCTGCTGATTTCATGTTACTTTCAGTACCTGTTTCATTGCGAGC 111
Qy 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 111
Db 254 TGTGTGGTGACTTACCAATTTATATGGACCACCCAGTAGAAGACTATATGAACCTAC 111
Qy 59 SerTyrHisSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 111
Db 314 ACATACCATTCAGTCTCACCTGCTTCAGTGAAGGAGCTATGCTGCAGAAAAATGCTGG 111
Qy 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheLeuSerSerGlu 111
Db 374 GGATGCTGCCCAATCACTGGAAGTCACTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 111
Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 111
Db 434 GAGAACTTCTGGAGCAGCAGTGAAGTGTGTTTCAGATGGGGCTCATCTGGTGGTG 111
Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 111
Db 494 ATCAATACTCAAGCGAGCAGAGATTTTCATCACCAGAGCTGAATGAGTCACTTCTTAC 111
Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 111
Db 554 TTCTCTGGGTCTTCGGATCCACAAAGTAATGGCAAAATGGCAATGGATGATGATCTCCT 111
Qy 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 111
Db 614 TTCAGTCAAAATGTGAGTGTCTGGCAGCCCATGAGCCCAATCTCCAGAGAGCGGTGT 111
Qy 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 111
Db 674 GTTTCATAGTTTACTGGAATCCTTCGAATGGGCTGGAGATGATGTTTCTGTGATAGT 111
Qy 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 734 AAACACAATTCAATATGTGAATGAAGAAGATTACCTA 772

RESULT 3
LOCUS BC023008 1254 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10, clone MGC:35901
IMAGE:5376247, mRNA, complete cds.
ACCESSION BC023008
VERSION BC023008.1 GI:18606437
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1254)
Strausberg, R.
Direct Submission
Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaphs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgeb@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 53 Row: p Column: 4  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9910157.

## FEATURES

## Location/Qualifiers

## source

1..1254

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="FVB/N-3"

/clone="MGC:35901 IMAGE:5376247"

/tissue\_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."

/clone\_lib="NCI-CGAP\_Mam2"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

146..775

/codon\_start=1

/product="C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10"

/protein\_id="AAH23008.1"

/db\_xref="GI:18606438"

/db\_xref="LocusID:56620"

/translation="MVQEROSQKGVCMTRLWSAAVISMILLSTCFIASCVVYQFI MDQPSRLYLHLTVINTEAEQNFITQOLNFSLSYFLGLSDPQNGKQWIDDDTPESQN SEQNCVOMGAHLVINTAEQNFITQOLNFSLSYFLGLSDPQNGKQWIDDDTPESQN VRFWHPHPNLPESRCVSIYWNPSKMGWNDVFCDSKHNSICEKKIYL"

BASE COUNT 362 a 271 c 276 g 345 t

ORIGIN

Alignment Scores:

Pred. No.: 3.04e-73 Length: 1254

Score: 793.00 Matches: 145

Percent Similarity: 77.00% Conservative: 19

Best Local Similarity: 68.08% Mismatches: 41

Query Match: 68.07% Indels: 8

DB: 10 Gaps: 4

US-09-766-511b-53 (1-209) x BC023008 (1-1254)

Qy 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18

Db 146 ATGGTGCAGAAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193

Qy 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSer 38

Db 194 AGACTCTGGTCAGCTGCTGTGATTTCCATGTTACTCTTTCAGTACCTGTTTCATTCGAGC 253

Qy 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58

Db 254 TGTGTGGTGACTTACCAATTTATTTATGGACCCAGCCAGTAGAGACTATATGAATTCAC 313

Qy 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76

Db 314 ACATACCATTCAGTCTCACCTGCTTCAGTGAAGGAGCTATGTTGTCAGAAAAATGTTGG 373

Qy 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96

Db 374 GGATGCTGCCAAATCACTGGAAGTCACTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 433

Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116

Db 434 GAGAACTTCTGGAGCAGCAGTGAAGTGTGTTTCAGATGGGGCTCATCTGGTGGTG 493

Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136

Db 494 ATCAATACTGAAGCGAGCAGAGATTTTCATCACCAGAGCTGAATGAGTCACTTCTTAC 553

Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156

Db 554 TTCTCTGGGTCTTCGGATCCACAAAGTAATGGCAAAATGGCAATGGATGATGATCTCCT 613



VERSION	AX155223.1	GI:14536702
KEYWORDS	human.	
SOURCE		
ORGANISM	human.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1312)	
AUTHORS	Schmitz,J., Dzionek,A. and Buck,D.W.	
TITLE	Antigen-binding fragments specific for dendritic cells, compositions and methods of use thereof antigens recognized thereby and cells obtained thereby	
JOURNAL	Patent: WO 0136487-A 1 25-MAY-2001; Milltenvi Biotec GmbH (DE)	

```

1. .1312
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="BDCA-2 cDNA sequence"
374 a 301 c 297 g 340 t
BASE COUNT
ORIGIN

```

Pred. No.:	6.22e-57	Length:	1312
------------	----------	---------	------

Percent Similarity:	68.08%
Best Local Similarity:	53.99%
Query Match:	54.68%
DB:	
Indels:	4
Gaps:	2
Mismatches:	64
Conservative:	30

US-09-766-511B-53 (1-209) x AX155223 (1-1312)

db 491 ATGGTACCTGAGCAAGAGCCTCAACACCGAGAGAAGGACCTGGTGGTTTCCAGTTGAAG 550

QY 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39

DB 551 GTCTGGTCCATGGGCAGTGGTATCCCATCTTGTCTCCACAGTGTCTGTTTCACTGTGAGTTCT 610

Db 611 GTGGTGCCTCACAAATTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGAGAG 670

QY 58 ---HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaIrrp 76  
||||| ||||||||| ||||| ::| |||

QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCys<sup>Tyr</sup>PheIleSerSerGlu 96

Db 731 AGCTGCTGCCCAACCCCTGGACTTCATTTCAGTCTAGTGGCTACTTTATTTCTACGGG 790

791 ATGCAATCTTGGCACTAAGAGCTCAAAGAAGCTGTTCTGTGATGGGGGGGATCTGGTGGTG 850

QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136

DU 001 ATCAGACCCAGGAGAAATGGAGATTTCATCATCTCAAGATCTCGAATAATGTTTCTTCTTAT 310

Ov 137 pheIeicIvIeiSerAsnDroGInCIvAsnAsnAsnTrnGInTrnTIIeaSnIvsThrPro 156

Db 911 TTTCTGGGGCTGTACAGATCCAGGGGGTTCGGCGACATTGGCAATGGGTTGACCAGACACCA 970

[illegible]

Qy 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196

Db 1031 GCGATAATAAAATTTCCGTTCTTCAGAAGAAATGGGGCTGGAATGACATTCTACTGTCACTA 1090

[illegible]

```

Db 1091 CCTCAGAAGTCAATTGCAAGATGAAGAAGATCTACATA 1129
RESULT 7
AF293615 1313 bp mRNA linear PRI 14-FEB-2002
LOCUS Homo sapiens blood dendritic cell antigen 2 protein (BDCA2) mRNA,
DEFINITION complete cds.
ACCESSION AF293615
VERSION AF293615.1 GI:17224597
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1313)
AUTHORS Dzionek,A., Sohna,Y., Nagafune,J., Cella,M., Colonna,M.,
Faccetti,F., Gunther,G., Johnston,I., Lanzavecchia,A.,
Nagasaka,T., Okada,T., Vermi,W., Winkels,G., Yamamoto,T., Zysk,M.,
Yamaguchi,Y. and Schmitz,J.
TITLE BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type
lectin, mediates antigen capture and is a potent inhibitor of
interferon alpha/beta induction
J. Exp. Med. 194 (12), 1823-1834 (2001)
JOURNAL 21615135
MEDLINE 11748283
PUBMED
REFERENCE 2 (bases 1 to 1313)
AUTHORS Sohna,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J.,
Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.
TITLE Cloning and characterization of blood dendritic cell antigen 2
(BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells
Unpublished
JOURNAL 3 (bases 1 to 1313)
AUTHORS Sohna,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J.,
Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-2000) Miltenyi Biotec GmbH, Friedrich-Ebert-Str.
68, Bergisch Gladbach 51429, Germany
FEATURES
source
1..1313
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="plasmacytoid dendritic cell; DC-2"
/gene="BDCA2"
492..1133
/gene="BDCA2"
/note="C-type lectin"
/codon_start=1
/product="blood dendritic cell antigen 2 protein"
/protein_id="AAL37036.1"
/db_xref="GI:17224598"
/translation="MVPPEEPQOREKGLWMFQKLVMSMAVVSILLLSVCFVTSVVPH
NFMYSKTVKRLSKREYQYHPSLTCVMEKDIEDWCCPTWTSFQSSCYFTSCMG
SWTKSKNCVMGADLVVINTREODFIQLNLRNSYFLGLSDPGRRHWQVDTF
YNEVTFWHSGEPNLDERCAILNFRSEEWGNDIHCHPQKSICKMKRIY"
BASE COUNT 374 a 302 c 297 g 340 t
ORIGIN
Alignment Scores:
Pred. No.: 6,22e-57 Length: 1313
Score: 637.00 Matches: 115
Percent Similarity: 68.08% Conservative: 30
Best Local Similarity: 53.99% Mismatches: 64
Query Match: 54.68% Indels: 4
DB: 9 Gaps: 2
US-09-766-511B-53 (1-209) x AF293615 (1-1313)
QY 1 MetMetGlnGlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArg 19
||||| ||||| ||||| ||||| |||||
Db 492 ATGGTGCCTGAAGAAGAGCCTCAAGACCGAGAGAAGGACTCTGGTGTCCAGTTGAAG 551
QY 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39

```

```

Db 552 GTCTGTCCATGCGCAGTCGTATCCATCTTGCCTCCTCAGTGTCTGTTCCTACTGTGAGTTCT 611
QY 40 ValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 612 GTGGTCCCTCAACAATTTATGTATACAAAAGTGTCAAGAGGCTGCCAAGTTACGAGAG 671
QY 58 ---HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaIleTrp 76
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 672 TATCAACAGTATCATCAAGCCCTGACCTCGTCATGCGAAGGAAAGACATAGAAGATTGG 731
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 732 AGCTGTGCGCCCAACCCCTTGGACTTCATTTCAGTCTAGTGTCTACTTTATTTCTACTGGG 791
QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
: : : : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 792 ATCAATCTTGGACTTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGGTG 851
QY 117 PheAsnThrGluAlaGlnGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 852 ATCAACACGAGGAGACACAGGATTTTCATCATTCACAATCTGAAAAAGAAATTTCTTCTTAT 911
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 912 TTCTCTGGGGCTGTTCAGATCCAGGGGTGGCGACATTTGGCAATGGGTTGACCAGACCA 971
QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
||| : : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 972 TACAATGAATAATTCGTTCTTCAGAGAATGGGGCTGGAATCACAATTCACATTCATCTGATGTA 1031
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
||| ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1032 CGCATAATAATTTCCGTTCTTCAGAGAATGGGGCTGGAATCACAATTCACATTCATCTGATGTA 1091
QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
: : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1092 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 1130
RESULT 8
AX357481 AX357481 827 bp DNA linear PAT 13-FEB-2002
LOCUS Sequence 19 from Patent WO0172773.
DEFINITION ACCESSION AX357481
ACCESSION AX357481
VERSION AX357481.1 GI:18674538
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.
TITLE Organic compounds
JOURNAL Patent: WO 0172773-A 19 04-OCT-2001;
Novartis AG (CH)
FEATURES
Source
1..827
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 242 a 164 c 188 g 233 t
ORIGIN
Alignment Scores:
Pred. No.: 1.92e-56 Length: 827
Score: 630.00 Matches: 114
Percent Similarity: 67.94% Conservative: 28
Best Local Similarity: 54.53% Mismatches: 63
Query Match: 54.08% Indels: 4
DB: 6 Gaps: 2
US-09-766-511B-53 (1-209) x AX357481 (1-827)
QY 5 GlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArgLeuTrpSerVal 23

```

```
Db 28 CAAGAGCCTCAAGACCCAGAGAAAGGACTCTGGTGGTTCCAGTTGAAGCTGGTCCATG 87
||||:||||| ||||| ||| |||:|||||
Qy 24 AlaclyIleSerIleAlaLeuSerAlaCysPheIleValSerCysValValThrTyr 43
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 GCAGTCGATCCATCTTGCTCCAGTGTCTGTTTCACTGTAGTCTCTGGTGGCTCAC 147
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 44 HisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSerTyr 60
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 148 AATTTATGATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGACAGATCAACAGTAT 207
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 CATCAAGCCTGACCTGCGTCATCGGAAGGAGACATAGAAGATGGAGCTGCTGCCCA 267
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 81 AlaserTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrp 100
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 ACCCTTGGACTTCATTTCACTGTAGTGTCTTATTTCTACTGGGATGCAATCTGG 327
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 ACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGTATCAACACG 387
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 388 GAAGACAGGATTCATCATTCAGAAATCTGAAAAGAAATTTCTTCTTATTTCTGGGCTG 447
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 TCAGATCCAGGGGTCGGCGACATCTGGCAATGGGTTTGACACACACATACATGAAAT 507
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal 180
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 GTACATCTCGGCACCTCAGGTGAACCAATACCTTGATGAGCGTTGTGGATAATAAT 567
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer 200
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 TTCGGTTCTTCAGAAAGTGGGCTGGAATGACATTCACGTCTGATGACCTCAGAAATCA 627
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 201 IleCysGluMetAsnLysIleTyrLeu 209
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 ATTGCAAGATGAAGAGATCTACATA 654
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AX357483 800 bp DNA linear PAT 13-FEB-2002
LOCUS
DEFINITION Sequence 21 from Patent WO0172773.
ACCESSION AX357483
VERSION AX357483.1 GI:18674539
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.
TITLE Organic Compounds
JOURNAL Patent: WO 0172773-A 21 04-OCT-2001;
Novartis AG (CH)
FEATURES
source Location/Qualifiers
1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 231 a 160 c 183 g 226 t
ORIGIN

Alignment Scores:
Pred. No.: 4.2e-55 Length: 800
Score: 617.00 Matches: 111
Percent Similarity: 67.14% Conservative: 30
Best Local Similarity: 52.86% Mismatches: 63
Query Match: 52.96% Indels: 6
DB: Gaps: 2
```

```
US-09-766-511B-53 (1-209) x AX357483 (1-800)
Qy 6 GlnProGlnSerThrGluLysArg-----GlyTrpLeuSerLeuArgLeuTrpSer 22
||||:||||| |||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 CCCCCCGGCTCGAAGACCGAGAGAAAGACTCTGGTGGTCCAGTTGCAAGGTCFGGTCC 61
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 23 ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr 42
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 ATGGCAGTCGATCATCTTCTCTCAGTGTCTCTTTCACTGTGAGTCTCTGGTGGCT 121
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 43 TyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSer 59
||||:||||| |||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 CACAATTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGAGAGTATCAACAG 181
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 60 TyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCys 79
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TATCATCAAGCTTCACTCGCTCATGGAAGGAAGACATAGAAGATTGGAGCTGTGCG 241
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 80 ProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluLysVal 99
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 CCAACCCCTTCGACTTCATTTCACTGTAGTGTCTTATTTCTACTGGGATGCAATCT 301
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 100 TrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThr 119
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TGGACTAAGAGTCAAAAGAACTGTTCTCTGTATGGGGCTGATCTGGTGTGATCAACACC 361
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 120 GluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGly 139
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 AGGGAAGAACAGAGATTTTCATTCATTCAGAAATCTGAAAGAAATTTCTTATTTCTGGGG 421
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 140 LeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLys 159
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 CTGTCAATGCCAGGGGTCGGCGACATTTGGCAATGGGTTGACCAGACCATCAATGAA 481
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 160 AsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIle 179
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 AATGTACATTTCTGCGACTCAGGTCAACCAATAACCTTGATGAGCGTTGTGCGATAATA 541
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 180 ValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsn 199
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 AATTTCCGTTCTTCAGAAAGATGGGGCTGGAATGACATTCATCTGTCATGTACCTCAGAAG 601
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 200 SerIleCysGluMetAsnLysIleTyrLeu 209
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 TCAATTTCAAGATGAAGAGATCTACATA 631
||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AF240359 1125 bp mRNA linear ROD 02-MAY-2000
LOCUS
DEFINITION Mus musculus strain BALB/c dectin-2 gamma isoform mRNA, complete
ACCESSION AF240359
VERSION AF240359.1 GI:7677474
KEYWORDS cds, alternatively spliced.
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1125)
Arizumi,K., Shen,G.-L., Shikano,S., Ritter,R. III, Zukas,P.,
Edelbaum,D., Morita,A. and Takashima,A.
Cloning of a second dendritic cell-associated C-type lectin
(dectin-2) and its alternatively spliced isoforms
J. Biol. Chem. (2000) in press
2 (bases 1 to 1125)
Arizumi,K. and Takashima,A.
Direct Submission
TITLE
JOURNAL Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical
REFERENCE Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA
AUTHORS Location/Qualifiers
1..1125
/organism="Mus musculus"
FEATURES
source
```

```

/strain="BALB/c"
/db_xref="taxon:10090"
/cell_line="X552"
/cell_type="dendritic cell"
146..652
/notes="C-type lectin; type II transmembrane protein;
dendritic cell-associated expression; C-type lectin motif;
alternatively spliced form of dectin-2 alpha mRNA"
/codon_start=1
/product="dectin-2 gamma isoform"
/protein_id="AAF67179.1"
/db_xref="GI:7677475"
/translation="MVQRSGQKGVCTLRLSAAVISMLLLSTCFIASCVVYQFI
MDQSRRLYELHYYLSLTCFSGTVMSEKMGCCPNHHKSGFSSCYLLISTKFNFWST
SEQNCVQMGAHLVINTEARQNFITQOLNESLYFLGLSNPSKWNVDVCDSKHNSI
CEMKKIYL"
BASE COUNT      320 a  243 c  249 g  313 t
ORIGIN
Alignment Scores:
Pred. No.:      2,29e-52      Length:      1125
Score:          592.50      Matches:      118
Percent Similarity: 62.91%      Mismatches: 16
Best Local Similarity: 55.40%      Conservative: 30
Query Match:      50.86%      Indels:      49
DB:              10      Gaps:      5

US-09-766-511B-53 (1-209) x AF240359 (1-1125)

QY  1  MetMetGlnGluGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 146 ATGTTGCAGGAAGACCAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCGTG 193

QY  19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuSerAlaCysPheIleValSer 38
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 194 AGACTCTGGTCAGCTGCTGTGATTTCCTCCATTTGAGTACCTGTTTCATTCGCAGC 253

QY  39 CysValValThrThrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 254 TGTCTGTGTGACTTACCAAAATTATTATGACGACGCCAGCCAGTATGAACTATATGAATTCAC 313

QY  59 SerTyrHisSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 314 ACATACCATTCACGTCACCTGCTTCAGTGAAGGACTATGGTCTCAGAAAAATGTGG 373

QY  77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 374 GGATGCTGCCAAATCACTGGAAGTCATTTGGCTCCAGCTGCTACCTCAITTCATCCAAAG 433

QY  97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 434 GAGAACTCTCTGGAGCACCACTGAGCAGAACTGTGTCAGATGGGGGCTCATCTGCTGGTG 493

QY  117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 494 ATCAATACTAAGCGAGCAGAAATTCATCACCAGCAGCTGAATGATGACTCACTTCTTAC 553

QY  137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 554 TTCCTGGGCTCTTCGAAT-----571

QY  157 TyrGluLysAsnValArgPheThrPheHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 571 -----571

QY  177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
    |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 572 -----CCTTCGAATGGGCTGGAATGATGATGTTTCTGTGATAGT 610

QY  197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
    ::: |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 611 AAACACAAATTCAATATCTGAATCTGAATGAAGAAGATTACCTA 649

```

RESULT 11	AF325460	739 bp	mRNA	linear	PRI 02-DEC-2001
LOCUS	AF325460				
DEFINITION	Homo sapiens dendritic lectin b isoform (CLECSF11) mRNA, complete cds, alternatively spliced.				
ACCESSION	AF325460				
VERSION	AF325460.1	GI:17225338			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Arce,I., Roda-Navarro,P., Montoya,M.C., Hernanz-Falcon,P., Puig-Kroger,A. and Fernandez-Ruiz,E.				
TITLE	Molecular and genomic characterization of human DLEC, a novel member of the C-type lectin receptor gene family preferentially expressed on monocyte-derived dendritic cells				
JOURNAL	Eur. J. Immunol. 31 (9), 2733-2740 (2001)				
MEDLINE	21426806				
PUBMED	11536172				
REFERENCE	2 (bases 1 to 739)				
AUTHORS	Arce,I., Hernanz,P. and Fernandez-Ruiz,E.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-DEC-2000) Biologia Molecular, Hospital de la Princesa, Diego de Leon, 62, Madrid 28006, Spain				
FEATURES	Location/Qualifiers				
source	1..739				
gene	/organism="Homo sapiens"				
CDS	/db_xref="taxon:9606"				
	/chromosome="12"				
	/map="12p12.3-p13.2"				
	1..739				
	/gene="CLECSF11"				
	8..556				
	/gene="CLECSF11"				
	/notes="C-type lectin receptor; DLEC-b; alternatively spliced"				
	/codon_start=1				
	/product="dendritic lectin b isoform"				
	/protein_id="AAL37359.1"				
	/db_xref="GI:17225339"				
	/translation="MVPEEPQDRVPHFMYSKVKRLSKLREYQQYHPSLTCVMEGK DIEDWCCPTPTSFQSCYFSTFGMSQTKSKGSMGADLVINTREQDFIION LKNSYVFLGLSDPGRRHWQVDQTPYENVTFWHSGPENLDERCAIINFRSSEEW GWNIDHCHVPQKSICKMKKIY"				
BASE COUNT	222 a	145 c	158 g	214 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1..57e-46	Length:	739		
Score:	534.50	Matches:	94		
Percent Similarity:	68.02%	Conservative:	23		
Best Local Similarity:	54.65%	Mismatches:	52		
Query Match:	45.88%	Indels:	3		
DB:	9	Gaps:	1		
US-09-766-511B-53 (1-209) x AF325460 (1-739)					
Qy	41	ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----	57		
Db	38	GTGCCTCACAAATTTATGTATAGCAAACTGTCAAGAGCGTCTCCAAAGTTACGAGACTAT	97		
Qy	58	HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGly	77		
Db	98	CAACAGTATCATCCAGCCCTACCTGCGTCATGAGGAAGACATAGAAATTGGAGC	157		
Qy	78	CysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheLeuSerSerGluGlu	97		
Db	158	TGCTGCCCAACCCCTTGAGACTTCATTTTCAGTCTAGTGTCTACTTTTCTACTGGGATG	217		
Qy	98	LysValTrpSerLysSerGluAlaAsnCysValGluMetGlyAlaHisLeuValPhe	117		



```

Db 218 CAATCTTGGACTAAGAGTCAAAGAAGCTGTTCTGTGTATGGGGCGTGATCTGGTGGTATC 277
Qy 118 AsnThrGluAlaGluInAsnPhelIeValGlnGlnLeuAsnGluSerPheSerTyrPhe 137
Db 278 AACACCGGGAGACAGGATTTCTATTCATTCAGAAATCTGAAAGAAATTTCTTATTTT 337
Qy 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
Db 338 CTGGGGCTGTTCAGATCCAGGGGTCGGCGACATTTGGCAATGGGTGACACAGACCATAC 397
Qy 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAla 177
Db 398 AATGAAATGTCATCTTGGCATCGAGTGAACCCCAATAACCTTGATGAGCGTGTGGC 457
Qy 178 SerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArg 197
Db 458 ATAATAATTTCCGTTCTTCAGAAGATGGGGCTGGGAATGACATTCATCTGTCATGTCCT 517
Qy 198 ArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 518 CAGAAGTCAATTTGCAAGATGAAGAATCTACATA 553

RESULT 12
LOCUS AX456975 444 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 14 from Patent WO0232958.
ACCESSION AX456975
VERSION AX456975.1 GI:21715770
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Smolyar, A.
TITLE Polynucleotide and polypeptide sequences of human dendritic cell
immunoreceptors
JOURNAL Patent: WO 0232958-A 14 25-APR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 123 a 86 c 101 g 134 t
ORIGIN

Alignment Scores:
Pred. No.: 5,75e-39 Length: 444
Score: 459.50 Matches: 80
Percent Similarity: 67.57% Conservative: 20
Best Local Similarity: 54.05% Mismatches: 47
Query Match: 39.44% Indels: 1
DB: 6 Gaps: 1

US-09-766-511B-53 (1-209) x AX456975 (1-444)
Qy 63 SerLeuThrCysPheSerGluGlyThrLysVal---ProAlaTrpGlyCysCysProAla 81
Db 1 TCTGTTAATAATTTTCTTCTTCTATATCTATTGGGAAGATTGGAGCTGCTCCCAACC 60
Qy 82 SerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrpSer 101
Db 61 CCTTGGACTTCATTCATCTAGCTAGTGTGCTACTTTTACTTTTCTACTGGGATGCAATCTGGACT 120
Qy 102 LysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAla 121
Db 121 AAGAGTCAAAAGAACTGTTCTGTGATGGGGGCTGATCTGGTGGTATCAACACAGGGAA 180
Qy 122 GluGlnAsnPhelIeValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSer 141
Db 181 GAACAGGATTTTCATTCATTCAGAAATCTGAAAGAAATTTCTTTATTTTCTGGGGCTGTCA 240
Qy 142 AspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnVal 161

```

```

Db 241 GATCCAGGGGTCGGCGACATTTGGCAATGGTTGACCACACACCATACATCAATGAAATGTC 300
Qy 162 ArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPhe 181
Db 301 ACATTTGGCACTCAGGTGAACCAATAACCTTGATGAGCGTGTGCGATAATAATTTTC 360
Qy 182 TrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIle 201
Db 361 CGTTCTTCAGAAGAATGGGGCTGGGAATGACATTCACCTGTCTACCTCAGAAGTCAATT 420
Qy 202 CysGluMetAsnLysIleTyrLeu 209
Db 421 TGCAAGATGAAGAAGATCTACATA 444

RESULT 13
LOCUS AX456962 402 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0232958.
ACCESSION AX456962
VERSION AX456962.1 GI:21715759
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Smolyar, A.
TITLE Polynucleotide and polypeptide sequences of human dendritic cell
immunoreceptors
JOURNAL Patent: WO 0232958-A 1 25-APR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..402
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 114 a 80 c 96 g 112 t
ORIGIN

Alignment Scores:
Pred. No.: 5,76e-39 Length: 402
Score: 459.00 Matches: 77
Percent Similarity: 70.90% Conservative: 18
Best Local Similarity: 57.46% Mismatches: 39
Query Match: 39.40% Indels: 0
DB: 6 Gaps: 0

US-09-766-511B-53 (1-209) x AX456962 (1-402)
Qy 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCystyrPheIleSerSer 95
Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTGTCTACTTTATTTCTACT 60
Qy 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
Db 61 GGGATGCAATCTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTG 120
Qy 116 ValPheAsnThrGluAlaGluGlnAsnPhelIeValGlnGlnLeuAsnGluSerPheSer 135
Db 121 GGTATCAACACACGGAAGAACAGGATTTTCATCTTCAGAAATCTGAAAAGAAATTTCTTCT 180
Qy 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155
Db 181 TATTTTCTGGGCTGTCAGATCCAGGGGTCGGCGACATTTGGCAATGGGTGACACAGACA 240
Qy 156 ProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln 175
Db 241 CCATCAATGAAATGTCACATTTCTGGCACTCAGGTGAACCCCAATAACCTTGATGAGCGT 300
Qy 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195
Db 301 TGTGGGATAATAAATTTCCGTTCTTCAGAAGAATGGGGCTGGGAATGACATTCACCTGTCTAT 360

```

```

QY 196 ThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
    ::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db 361 GTACCTCAGAGTCAATTGCAAGATGAAGAAGATCTACATA 402

RESULT 14
AF109146
LOCUS
DEFINITION
Homo sapiens C-type lectin superfamily 6 (CLECSF6) mRNA, complete
cds.
ACCESSION
AF109146
VERSION
AF109146.1 GI:6502534
KEYWORDS
Homo sapiens.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1133)
AUTHORS
Richard, M. and Beaulieu, A.D.
TITLE
A novel C-type lectin expressed in GM-CSF stimulated neutrophils
JOURNAL
Unpublished
AUTHORS
Richard, M. and Beaulieu, A.D.
TITLE
Direct Submission
JOURNAL
Submitted (24-NOV-1998) Dept. of Medicine, Centre Hospitalier de
1'Universite Laval, 2705 Boul. Laurier, Sainte-Foy, Quebec G1V 4G2,
Canada
FEATURES
source
1..1133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="polymorphonuclear cells"
1..1133
/gene="CLECSF6"
95..808
/gene="CLECSF6"
/feature="C-type lectin DBD27"
/codon_start=1
/product="C-type lectin superfamily 6"
/protein_id="AAFI4348.1"
/db_xref="GI:6502535"
/translation="MTSEITVAEVRFKNEFKSSGINTASSAASKERTAPLKSNTGPPK
LLCSLLIFFLLAIISFFIAFVIFQKYSQLEKKTKKELVHTLVCVKKNMPVEETA
WSCCPKNKWSFSSNCYFISTESASQDSEKDKARMEAHLLVINTOEQDFIFONLQEE
SAYFVGLSDPEGRHWQWVDQTPYNESSTFWHPREPSDPNERCVCVNLFRKSPKRWGN
DVNCLGPQRSVCMMKHL"
BASE COUNT 332 a 214 c 239 g 348 t
ORIGIN
Alignment Scores:
Pred. No.: 2.54e-38 Length: 1133
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 9 Gaps: 3

US-09-766-511b-53 (1-209) x AF109146 (1-1133)
QY 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
    ||| |||||::: ||| |||||::: ||| |||||::: ||| |||||:::
Db 248 ATATTTTCTCCTGCTATGCAATCTCATCTTTATTTGCTTTTGTCTATTTCTTTCAA--- 304

QY 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerLeuThr 65
    ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 305 AAATATCTCAGCTTCTTCAAAAAGAGCTACAAAAGAGCTGGTTTCATACACATTGGAG 364

QY 66 CysPheSerGluGlyThrLysVal-----ProAlaThrGlyCysCysProAlaSerTrp 83
    ||| ::| ::| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 TGTGTGAAAAAATAATGCGCCGTGGAAGACAGAGCTGTGGCCCAAGAATTCG 424

QY 84 LysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 AAGTCATTAGTTCCAACTGCTACTTATTCTACTGATCAGCATCTTGGCAAGACAGT 484

```

```

QY 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123
    |||||::: ||| |||||::: ||| |||||::: ||| |||||::: |||||:::
Db 485 GAGAAGGACTGTGCTAGAATGGAGGCTCAGCTGCTGATAAAGACTCAAGAAGAGCAG 544

QY 124 AspPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
    ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 GATTTCATCTCCAGATCTGCGAAGAGAAATCTGCTATTATTTTGGGGCTCTCAGATCCA 604

QY 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
    ::| ||| ::| |||||::: |||||::: |||||::: |||||::: |||||:::
Db 605 GAAGGTCAAGCACATTCGCAATGGTTGATCAGACACCATACAAATGAAGATTCACATTC 664

QY 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
    ||||| |||||::: |||||::: |||||::: |||||::: |||||:::
Db 665 TGGCATCCAGCTCAGGCCAGCTGATCCCAATGAGCGCTGCTGCTAAATTTTCGTAAA 724

QY 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 725 TCACCCAAAAGATGGGGCTGGAATGATGTTAATTGCTTGGTCTCAAAGGTCAGTTGT 784

QY 203 GluMetAsnLysIleTyrLeu 209
    ||||| |||||::: |||||::: |||||::: |||||::: |||||:::
Db 785 GAGATGATGAGATCCACITTA 805

RESULT 15
HSAL33532
LOCUS
DEFINITION
Homo sapiens mRNA for dendritic cell immunoreceptor.
ACCESSION
AJ133532
VERSION
AJ133532.1 GI:5823973
KEYWORDS
DCIR gene; dendritic cell immunoreceptor.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1271)
AUTHORS
Bates, E.E., Fournier, N., Garcia, E., Valladeau, J., Durand, I.,
Pin, J.J., Zurawski, S.M., Patel, S., Abrams, J.S., Lebecque, S.,
Garrone, P. and Saeland, S.
TITLE
APCs express DCIR, a novel C-type lectin surface receptor
containing an immunoreceptor tyrosine-based inhibitory motif
JOURNAL
J. Immunol. 163 (4), 1973-1983 (1999)
MEDLINE
99370001
PUBMED
10438934
REFERENCE
2 (bases 1 to 1271)
AUTHORS
Bates, E.E.
TITLE
Direct Submission
JOURNAL
Submitted (09-MAR-1999) Bates E.E., Laboratory for Immunological
Research, Schering-Plough, 27 Chemin des Peupliers, BP11, 69571
Bardilly, FRANCE
COMMENT
chromosomal localisation by RH mapping 12p13.
FEATURES
source
1..1271
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="dendritic"
242..955
/gene="DCIR"
242..955
/gene="DCIR"
/function="cell surface receptor"
/codon_start=1
/product="dendritic cell immunoreceptor"
/protein_id="CAB54001.1"
/db_xref="GI:5823974"
/translation="MTSEITVAEVRFKNEFKSSGINTASSAASKERTAPIKSNTPGPK
LUCASLLIFFLLAIISFFIAFVIFQKYSQLEKKTKKELVHTLVCVKKNMPVEETA
WSCCPKNKWSFSSNCYFISTESASQDSEKDKARMEAHLLVINTOEQDFIFONLQEE
SAYFVGLSDPEGRHWQWVDQTPYNESSTFWHPREPSDPNERCVCVNLFRKSPKRWGN
DVNCLGPQRSVCMMKHL"
BASE COUNT 349 a 243 c 279 g 400 t
ORIGIN

```

```
Alignment Scores:
Pred. NO.:      2.92e-38      Length:      1271
Score:          458.00      Matches:      83
Percent Similarity: 62.57%      Conservative: 34
Best Local Similarity: 44.39%      Mismatches: 66
Query Match:      39.31%      Indels:      4
DB:                9      Gaps:        3

US-09-766-511B-53 (1-209) x HSA133532 (1-1271)
Qy 26 IleSerIleAlaLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
   ||| |||||:|:| ||| :|:|:| |||:|:|:|
Db 395 ATATTTCCTCGTATTGGCAATCTCATCTTATTGCTTTTGTCTTTCTTCAA--- 451
Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
   ||| |||:|:|:|:|:| |||:|:|:|:|:| |||:|:|:|:|:|
Db 452 AAATATTCTCAGCTCTTGAAGAAAAGACTACAAAGAGCGTGGTTTCATACACATTGGAG 511
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83
   ||| :|:| ||| ||||| ||||| ||||| ||||| |||||
Db 512 TGTGTGAAGAAATATGCCCGTGGAGAGACAGACCTGGAGCTGTGCCCAAGAAATTGG 571
Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGlyLysValTrpSerLysSer 103
   ||||| |||:|:|:|:|:| ||||| ||||| ||||| ||||| |||||
Db 572 AAGTCATTAGTTCCAACTGCTACTTATTCTTACTGAATCAGCATCTTGGCAAGACAGT 631
Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123
   |||:|:|:|:|:| ||| ||| ||||| ||||| ||||| ||||| |||||
Db 632 GAGAGGACTGTGCTAGAAATGGAGGCTCACCTGCTGGTATAACACTCAAGAGAGAGCAG 691
Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
   :|:| ||||| ||| ||| ||| :|:| |||||:|:| ||||| ||||| |||||
Db 692 GATTTTCATCTCCAGATCTGCAAGAGAAATCTGCTTATTTTGCGGGCTCTCAGATCCA 751
Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
   :|:| ||| :|:| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 GAAGGTCAGCGACATTGGCAATGGTTGATCAGACACCACATACAATGAAAGTCCACATTC 811
Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
   ||||| ||||| |||||:|:| |||:|:| |||:|:| |||:|:| |||
Db 812 TGGCATCCAGTGGAGCCCGCTGATGCCAATGAGCGCTGGCTGTGCTAAATTTTCGTAAA 871
Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 872 TCACCCAAAGATGGGCTGGAATGATGTTAATTGTCTTGGTCTCTCAAGGTCAGTTTGT 931
Qy 203 GluMetAsnLysIleTyrLeu 209
   ||||| ||||| |||||:|:|
Db 932 GAGATGATGAAGATCCACTTA 952
```

Search completed: February 20, 2003, 20:11:03  
Job time : 2236 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 19:29:57 : Search time 58 Seconds  
(without alignments)  
1835.321 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQOPSTKRGWLSRL.....NDVICETRNICEWNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 424239 segs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094339\_29351/app\_query.fasta\_1.391  
-DB=PublishedApplications.NA -QPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09766511\_cgn\_1\_1\_28\_erunat\_19022003\_094339\_29351  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	765.5	65.7	693	10	US-09-833-381-994 Sequence 994, Appl
C 2	653.5	56.1	528	10	US-09-833-381-995 Sequence 995, Appl
3	637	54.7	642	12	US-10-090-466-1 Sequence 1, Appl
4	535.5	46.0	549	12	US-10-090-466-3 Sequence 3, Appl

5	458	39.3	1091	10	US-09-764-870-199 Sequence 199, App
6	458	39.3	1096	10	US-09-764-870-38 Sequence 38, Appl
7	458	39.3	1104	9	US-09-862-802-1 Sequence 1, Appl
C 8	457.5	39.3	1036	10	US-09-833-381-119 Sequence 119, App
9	440.5	37.8	758	10	US-09-833-381-1339 Sequence 1339, Ap
10	436	37.4	1418	9	US-09-862-802-7 Sequence 7, Appl
11	398.5	34.2	968	9	US-09-965-529-40 Sequence 40, Appl
12	398.5	34.2	997	9	US-09-905-291A-376 Sequence 376, App
13	398.5	34.2	997	9	US-09-902-853-376 Sequence 376, App
14	398.5	34.2	997	9	US-09-907-824-376 Sequence 376, App
15	398.5	34.2	997	9	US-09-907-841-376 Sequence 376, App
16	398.5	34.2	997	9	US-09-904-011-376 Sequence 376, App
17	398.5	34.2	997	9	US-10-174-590-23 Sequence 23, Appl
18	398.5	34.2	997	9	US-10-176-758-23 Sequence 23, Appl
19	398.5	34.2	997	9	US-10-175-737-23 Sequence 23, Appl
20	398.5	34.2	997	9	US-09-906-742-376 Sequence 376, App
21	398.5	34.2	997	9	US-10-173-708-23 Sequence 23, Appl
22	398.5	34.2	997	9	US-10-175-738-23 Sequence 23, Appl
23	398.5	34.2	997	9	US-10-175-752-23 Sequence 23, Appl
24	398.5	34.2	997	9	US-10-176-482-23 Sequence 23, Appl
25	398.5	34.2	997	9	US-10-176-757-23 Sequence 23, Appl
26	398.5	34.2	997	9	US-10-176-913-23 Sequence 23, Appl
27	398.5	34.2	997	9	US-10-180-552-23 Sequence 23, Appl
28	398.5	34.2	997	9	US-10-180-557-23 Sequence 23, Appl
29	398.5	34.2	997	9	US-09-906-838-376 Sequence 376, App
30	398.5	34.2	997	9	US-09-907-613-376 Sequence 376, App
31	398.5	34.2	997	9	US-09-907-942-376 Sequence 376, App
32	398.5	34.2	997	9	US-10-173-700-23 Sequence 23, Appl
33	398.5	34.2	997	9	US-10-174-572-23 Sequence 23, Appl
34	398.5	34.2	997	9	US-10-174-579-23 Sequence 23, Appl
35	398.5	34.2	997	9	US-10-174-582-23 Sequence 23, Appl
36	398.5	34.2	997	9	US-10-174-588-23 Sequence 23, Appl
37	398.5	34.2	997	9	US-10-175-739-23 Sequence 23, Appl
38	398.5	34.2	997	9	US-10-175-740-23 Sequence 23, Appl
39	398.5	34.2	997	9	US-10-175-743-23 Sequence 23, Appl
40	398.5	34.2	997	9	US-10-176-488-23 Sequence 23, Appl
41	398.5	34.2	997	9	US-10-176-492-23 Sequence 23, Appl
42	398.5	34.2	997	9	US-10-176-747-23 Sequence 23, Appl
43	398.5	34.2	997	9	US-10-176-750-23 Sequence 23, Appl
44	398.5	34.2	997	9	US-10-176-985-23 Sequence 23, Appl
45	398.5	34.2	997	9	US-10-176-987-23 Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-09-833-381-994/c  
; Sequence 994, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 994  
; LENGTH: 693  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(693)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-994

Alignment Scores:  
Pred. No.: 5e-78 Length: 693  
Score: 765.50 Matches: 138  
Percent Similarity: 78.28% Conservative: 17

Best Local Similarity: 69.70% Mismatches: 39  
Query Match: 65.71% Indels: 4  
DB: 10 Gaps: 2  
US-09-766-511b-53 (1-209) x US-09-833-381-994 (1-693)

QY 15 TrpLeuSerLeuArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCys 34  
DB 674 TGG--ACCTGAGACTCTGGTCAGCTGCTGTGATTTCCATTTACTTCTTGAGTACCTGT 618  
QY 35 PheIleValSerCysValValThrTyrHisPheThrTyrGlyGluThrGlyLeu 54  
DB 617 TTCATTGCCAGCTGTGTGTGACTTACCAATTTATATGGACCACCCAGTAGAAGACTA 558  
QY 55 SerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal--- 73  
DB 557 TATGAACCTTCACACATACCATTCCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGCA 498  
QY 74 ---ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPhe 92  
DB 497 GAAAAAATGTGGGATGCTGCCCAATCACCTGGAAGTCATTTGGCTCCAGCTGTACCTC 438  
QY 93 IleSerSerGluGluLysVal-TripSerLysSerGluGlnAsnCysValGluMetGlyAl 112  
DB 437 ATTCTACCAAGAGAACTTCTGGAGCACCACTGAGCAGAACTGTTCATATGGGGGC 378  
QY 112 aHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnLeuAsnG 132  
DB 377 TCATCTGGTGGTGATCAATACTGAAGCGAGCAGAGAATTTTCATCACCAGCAGCTGAATGA 318  
QY 132 uSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIl 152  
DB 317 GTCACTTTNTACTTCTGGGTCCTTTCGGATCCACAAAGGTAATGGCAATGGAT 258  
QY 152 eAspLysThrProTyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSe 172  
DB 257 CGATGATACCTCTTCAGTCAAAATGTGAGTCTGGCACCCCATGAACCAATCTTCC 198  
QY 172 rAlaGluGlnCysAlaSerIleValPheThrLysProThrGlyTrpGlyTrpAsnAspVa 192  
DB 197 AGAAGAGCGGTGTGTTCATAGTTTACTTGGAAATCCTTCGAAATGGGGCTGGAATGATG 138  
QY 192 lIleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
DB 137 TTTCTGTGATGAACACAAATTCATATGTGAATGAAGAAGATTTACCTA 86

RESULT 2  
US-09-833-381-995/C  
; Sequence 995, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 995  
; LENGTH: 528  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(528)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-995

Alignment Scores:  
Pred. No.: 1.78e-65 Length: 528  
Score: 653.50 Matches: 120

Percent Similarity: 78.11% Conservative: 12  
Best Local Similarity: 71.01% Mismatches: 34  
Query Match: 56.09% Indels: 3  
DB: 10 Gaps: 2  
US-09-766-511b-53 (1-209) x US-09-833-381-995 (1-528)  
QY 15 TrpLeuSerLeuArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCys 34  
DB 509 TGG--ACCTGAGACTCTGTGTCAGCTGCTGTGATTTCCATTTACTTCTTGAGTACCTGT 453  
QY 35 PheIleValSerCysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeu 54  
DB 452 TTCATTGCCAGCTGTGTGTGACTTACCAATTTATATGGACCACCCAGTAGAAGACTA 393  
QY 55 SerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal--- 73  
DB 392 TATGAACCTTCACACATACCATTCCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGCA 333  
QY 74 ---ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPhe 92  
DB 332 GAAAAAATGTGGGATGCTGCCCAATCACCTGGAAGTCATTTGGCTCCAGCTGTACCTC 273  
QY 93 IleSerSerGluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAla 112  
DB 272 ATTCTACCAAGAGAACTTCTGGAGCACCACTGAGCAGAACTGTGTTCAGATGGGGCT 213  
QY 113 HisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnLeuAsnGlu 132  
DB 212 CATCTGGTGGTGATCAATACTGAAGCGGAGCAGAAATTTTCATCACCAGCAGCTGAATGAG 153  
QY 133 SerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIle 152  
DB 152 TCATCTTTTACTTCTCTGGGCTCTTCGGATCCACAAAGTAAATGGCAATGGCAATGATC 93  
QY 153 AspLysThrProTyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSer 172  
DB 92 GATGATACCTCTTTCAGTCAAAATGTGAGTCTTGGCACCCCATGAACCAATCTTCCA 33  
QY 173 AlaGluGlnCysAlaSerIleValPhe 181  
DB 32 GAAGACGGGTGTGTTCAATAGATTAC 6

Alignment Scores:  
Pred. No.: 1.79e-63 Length: 642  
Score: 637.00 Matches: 115  
Percent Similarity: 68.08% Conservative: 30  
Best Local Similarity: 53.99% Mismatches: 64  
Query Match: 54.68% Indels: 4  
DB: 12 Gaps: 2

US-09-766-511B-53 (1-209) x US-10-090-466-1 (1-642)

```
QY 1 MetMetGlnGluGlnProGlnSerThrGluLys--ArgGlyThrLeuSerLeuArg 19
||||: |||:||||| ||||| ||||| ||||| |||||
Db 1 ATGGTGCTGAAGAAGAGCCTCAAGACCGAGAGAAGGACTGGTGGTTCAGATTGAAG 60

QY 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39
:::|||||:||||| |||:||||| ||||| ||||| ||||| |||||
Db 61 GTCTGGTCCATGGCAGTCGATCCATCTGCTCCTCAGTGTCTTCTCACTGTGAGTTCT 120

QY 40 ValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
|||||:||||| |||:||||| ||||| ||||| ||||| ||||| |||||
Db 121 GTGGTGCTCACAATTTTATGTATACAAAACTGTCAAGAGGCTGTCCAAAGTTACGAGAG 180

QY 58 ---HisSerTyrHisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TATCAACAGTATCATCAAGCCTGACCTGGTCAATGGAAGGAAGACATAGAAGATTGG 240

QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGlu 96
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGCTGTGCCCAACCCCTTGGACTTCATTTCACTCAGTCTAGTTGCTACTTTATTCTACTGGG 300

QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
::: |||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ATCAATCTTGGACTAAGAGTCAAAAGAACTGTCTGTGTATGGGGCTGATCTGGTGGT 360

QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 ATCAACACCGAGGAAGACAGGATTTTCATCTCAATTCAGAAATCTTCTTCTTAT 420

QY 137 PheLeuGlyLeuSerAspProGlnGlnAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 TTCTCTGGGGCTGTCAGATCCAGGGGCTCGGCGACATTGGCAATGGGTTTCAACACACCA 480

QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 TACAATGAAATGTCACTTCTGGCAGTCAAGTGAACCCCAATACCTTGATGAGCGTTGT 540

QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 GCGATAATAATTTCCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACCTGTCATGTA 600

QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 639
```

## RESULT 4

```
; Sequence 3, Application US/10090466
; Patent No. US20020137914A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mather, Brian
; APPLICANT: Cullinan, Emily B.
; TITLE OF INVENTION: No. US20020137914A1el Human Dectin Proteins and Polynucleotides B
; FILE REFERENCE: LEX-0315-USA
; CURRENT APPLICATION NUMBER: US/10/090,466
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/274,961
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-090-466-3
```

Alignment Scores: 4.86e-52 Length: 549  
Pred. No.: 535.50 Matches: 94  
Score: 68.02% Conservative: 23  
Percent Similarity: 54.65% Mismatches: 52

Query Match: 45.97% Indels: 3  
DB: 12 Gaps: 1

US-09-766-511B-53 (1-209) x US-10-090-466-3 (1-549)

```
QY 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
||||: |||:||||| |||:||||| ||||| ||||| ||||| |||||
Db 31 GTGCTCACAATTTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAGTTACGAGAGTAT 90

QY 58 HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGly 77
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 CAACAGTATCATYCAAGCCTGACCTGCCTCATGGAAGGAAGGACATAGAAGATTGGAGC 150

QY 78 CysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGlu 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 TGCTGCCCAACCCCTTGGACTTCATTTCACTAGTCTACTTTATTTCTACTGGGATG 210

QY 98 LysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPhe 117
::: |||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 CAATCTTGGACTAAGAGTCAAAAGAACTGTCTGTGTATGGGGCTGATCTGTGTGTGATC 270

QY 118 AsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPhe 137
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 AACACGAGGAAGAAGAGGATTTTCATCTCAGATCTGAAAAGAAATCTCTTATTTT 330

QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 CTGGGGCTGTCAAGTCCAGGGGCTCGGCGACATTTGGCAATGGTGGTACCACACCATAC 390

QY 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAla 177
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 AATGAAATGTCACTTCTGGCAGTCAAGTGAACCAATAACCTTGATGAGCGTTGTGCG 450

QY 178 SerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArg 197
||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 ATAATAAATTTCCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACCTGTCATGTACCT 510

QY 198 ArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 CAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 546
```

## RESULT 5

```
US-09-764-870-199
; Sequence 199, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 1091
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-199
```

Alignment Scores: 8.75e-43 Length: 1091  
Pred. No.: 458.00 Matches: 83  
Score: 62.57% Conservative: 34  
Percent Similarity: 44.39% Mismatches: 66  
Best Local Similarity: 39.31% Indels: 4  
Query Match: 10 Gaps: 3

US-09-766-511B-53 (1-209) x US-09-764-870-199 (1-1091)

```
QY 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 ATATTTTCTGCTATTGGCAATCTCATCTTTATTTGCTTTGTCATTTCTTTCAA--- 228
```

```
Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 AAATATTCTCAGCTTCTTGAAAGAAAGACTACAAAAGAGCTGGTTTCATACAACATTGGAG 288

Qy 56 CysPheSerGluGlyThrLysVal-----ProLalaTrpGlyCysCysProAlaSerTrp 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 TGTGTGAAAAAATAATGCCCGTGAAGAGACAGCCTGGAGCTGTTGCCCAAAGAATTGG 348

Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGlyLysValTrpSerLysSer 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 AAGTCATTAGTTCCTCACTGCTTATTCTACTGAATCAGCATCTTGGCAAGACAGT 408

Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGluGln 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GAGAAGGACTGTGTAGAAATGGAGGCTCACCTGCTGGTGATAAACACTCAAGAAGACAG 468

Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 GATTTTCATCTCCAGAATCTCAAGAAGAATCTGCTTATTATTGTTGGGCTCTCAGATCCA 528

Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 GAAGGTCAGCGACATGCGCAATGGTTGATCAGACACACCATACATGAAGATTCACATTC 588

Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 TGGCATCCACGTGAGCCAGTGATCCCAATGAGCGCTGCGTTGGCTAAATTTTCGTAAA 648

Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 TCACCCAAAAGATGGGCTGGAATGATGTTAATTGCTTGTCTCCTCAAGGTCAGTTTGT 708

Qy 203 GluMetAsnLysIleTyrLeu 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATGATGAAGATCCACTTA 729

RESULT 6
US-09-764-870-38
; Sequence 38, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-38

Alignment Scores:
Pred. No.: 8,81e-43 Length: 1096
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 10 Gaps: 3

US-09-766-511b-53 (1-209) x US-09-764-870-38 (1-1096)

Qy 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 ATATTTTCCCTGCTATTGGCAATCTCATCTTATTGCTTTTGCATCTTTCTTTCAA-- 238

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 AAATATTCTCAGCTTCTTGAAAGAAAGACTACAAAAGAGCTGGTTTCATACAACATTGGAG 298
```

```
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 TGTGTGAAAAAATAATGCCCGTGAAGAGACAGCCTGGAGCTGTGCCCCAAAGAATTGG 358

Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGlyLysValTrpSerLysSer 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 AAGTCATTAGTTCCTCACTGCTTATTCTACTGAATCAGCATCTTGGCAAGACAGT 418

Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGluGln 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 GAGAAGGACTGTGTAGAAATGGAGGCTCACCTGCTGGTGATAAACACTCAAGAAGACAG 478

Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 GATTTTCATCTCCAGAATCTGCAAGAAGAATCTGCTTATTATTGTTGGGCTCTCAGATCCA 538

Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 GAAGGTCAGCGACATGCGCAATGGTTGATCAGACACCATACATGAAGATTCACATTC 598

Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 TGGCATCCACGTGAGCCAGTGATCCCAATGAGCGCTGCGTTGCTAAATTTTCGTAAA 658

Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 TCACCCAAAAGATGGGCTGGAATGATGTTAATTGCTTGTGCTCCTCAAGGTCAGTTTGT 718

Qy 203 GluMetAsnLysIleTyrLeu 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 GAGATGATGAAGATCCACTTA 739

RESULT 7
US-09-862-802-1
; Sequence 1, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 242...952
; OTHER INFORMATION: protein coding sequence
US-09-862-802-1

Alignment Scores:
Pred. No.: 8,9e-43 Length: 1104
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 9 Gaps: 3

US-09-766-511b-53 (1-209) x US-09-862-802-1 (1-1104)

Qy 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ATATTTTCCCTGCTATTGGCAATCTCATCTTATTGCTTTTGTCTATTTCCTTCAA--- 451

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
```



US-09-766-511B-53 (1-209) x US-09-833-381-1119 (1-1036)

Qy 29 AlaLeuLeuSerAlaCysPheIleValSerCysValValThrThyHisPheThrThyGly 48  
          :::|||     ::: |||       ::: |||:::

Db 701 AGCCCTCAGGACCGAGTGCCTCACAAATTTATGTATG----- 666

Qy	42	ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr
Db	69	ATTGCTGAGTGTTCATCTTCTCAGGTCTGTTTTATTCCAGTGTGTTGGTACT
Qy	61	TyrHis---PheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHis
Db	129	CATCACAACTTTTTCACCGCTGTAAAGAGAGACACAGAGTGCACAAAGTTA



Alignment Scores:  
Pred. No.: 4,23e-36 Length: 968  
Score: 398.50 Matches: 84  
Percent Similarity: 57.28% Conservative: 34  
Best Local Similarity: 40.78% Mismatches: 81  
Query Match: 34.21% Indels: 7  
DB: Gaps: 5

US-09-766-511B-53 (1-209) x US-09-965-529-40 (1-968)

Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20  
Db 174 AATCATCTGAACACAAATGCACAGAG---AGAGGATGCTCTCTCCCAAAATGTTCTTA 230  
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
Db 231 TGGACTGTCGGATGCCCATCTATTTCTCAGTCGCTGTTTCTCATCCAGATGTTGTT 290  
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 291 GTGACATTCGCATC---TTTCAAACTGTGATGAGAAAAGTTTCAGCTACCTGAGAAT 347  
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 348 TTCACAGAGCTCTCTCTCTACAAATATGATCAGGT---TCAGTCAAGAAATTTCTGTCCA 404  
Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
Db 405 TTGAACCTGGGAATATTTTCAATCCAGCTGCTACTTCTTTCTACTGACACCATTTCTCTGG 464  
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
Db 465 CGGTAACTTAAAGAACTGCTAGCCATGGGGCTCACCCTGGTGTATCACTACACAG 524  
Qy 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
Db 525 GAGGAGCAGGAATTCCTTCTTCAAGAAACCTAAATGAGAGAGTTTTTATTGGAGCTG 584  
Qy 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
Db 585 TCAGACCAAGGTGTCGAGGGTCAGTGGCAATGGGTGGCGGACACACCTTTGACAAAGTCT 644  
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178  
Db 645 CTGAGCTCTCGGATGTAGGGAGGCCCAACAACATAGCTACCTGGAGGACTGTGCCACC 704  
Qy 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198  
Db 705 ATGAGAGACTCTTCAAAACCAAGGCAAAATTTGAATGATGATTAACCTGTTTCTCAATTAT 764  
Qy 199 AsnSerIleCysGluMet 204  
Db 765 TTTCGGATTGTGAAATG 782

RESULT 12

US-09-905-291A-376  
Sequence 376, Application US/09905291A

Patent No. US20020160374A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: ROY, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/905/291A  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 376  
LENGTH: 997  
TYPE: DNA  
ORGANISM: Homo Saplen  
US-09-905-291A-376

Alignment Scores:  
Pred. No.: 4,42e-36 Length: 997  
Score: 398.50 Matches: 84  
Percent Similarity: 57.28% Conservative: 34  
Best Local Similarity: 40.78% Mismatches: 81  
Query Match: 34.21% Indels: 7  
DB: Gaps: 5

US-09-766-511B-53 (1-209) x US-09-905-291A-376 (1-997)

Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20  
Db 118 AATCATCTGAACACAAATGCACAGAG---AGAGGATGCTCTCTCCCAAAATGTTCTTA 174  
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40

175	TGGACTGTTGCTGGGATCCCAATCTCTATTTCTCAGTCCCTGTTTTCATCACCAGATGTGTT	234	
Qy	41	ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr	60
Db	235	GTGACATTTCCGATC---TTTCAAACTGTGATGACAAAAGTTTCAGTACTGAGAA	291
Qy	61	HisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro	80
Db	292	TTCACAGAGCTCTCTGCTACAATTATGGATCAGGT---TCAGTCAAGAAATTTGTTCTCA	348
Qy	81	AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp	100
Db	349	TTGAACGTGGGAATATTTTCAATCCAGCTGCTACTCTTTTCTACTGACACCATTTCTCTGG	408
Qy	101	SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGlu	120
Db	409	CGGTTAAGTTTAAGAACTGCTCAGGCATGGGGGCTACACCTGGTGTATTCAACTCACAG	468
Qy	121	AlaGluGlnAsnPheIleValGlnGlnLeuAsnGlnSerPheSerTyrPheLeuGlyLeu	140
Db	469	GAGGAGCAGGAATTCCTTCTTCAAGAAACCTAAATCAGAGAGAGTTTATTTCGAGCTG	528
Qy	141	SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn	160
Db	529	TCAGACCAAGTTCTCGAGGGTCAGTGGCAATGGGTGGACGGCACACCTTTGACAAAGTCT	588
Qy	161	ValArgPheThrHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer	178
Db	589	CTGAGCTTCGGGATGTAGGGAGCCCAACAATAGCTACCCCTGGAGGAGCTGTGCCACC	648
Qy	179	IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg	198
Db	649	ATCAGAGAGCTTTCAAACCCAGGCAAAATTCGAATGATGTAACTGTTTCTCTCAATTAT	708
Qy	199	AsnSerIleCysGluMet	204
Db	709	TTTCGGGATTTGTGAAATG	726

```

RESULT 13
US - 09-902-853-376
  / Sequence 376, Application US/0902853
  / Publication No. US20020192659A1
  / GENERAL INFORMATION:
  / APPLICANT: Genentech, Inc.
  / APPLICANT: Ashkenazi, Avi
  / APPLICANT: Botstein, David
  / APPLICANT: Desnoyers, Luc
  / APPLICANT: Eaton, Dan L.
  / APPLICANT: Ferrara, Napoleone
  / APPLICANT: Filvaroff, Ellen
  / APPLICANT: Fong, Sherman
  / APPLICANT: Gao, Wei-Qiang
  / APPLICANT: Gerber, Hanspeter
  / APPLICANT: Gerritsen, Mary E.
  / APPLICANT: Goddard, A.
  / APPLICANT: Godowski, Paul J.
  / APPLICANT: Grimaldi, Christopher J.
  / APPLICANT: Gurney, Austin L.
  / APPLICANT: Hillan, Kenneth, J.
  / APPLICANT: Klijavin, Ivar J.
  / APPLICANT: Mather, Jennie P.
  / APPLICANT: Pan, James
  / APPLICANT: Paoni, Nicholas F.
  / APPLICANT: Ray, Margaret Ann
  / APPLICANT: Stewart, Timothy A.
  / APPLICANT: Tumas, Daniel
  / APPLICANT: Williams, P. Mickey
  / APPLICANT: Wood, William, I.
  / TITLE OF INVENTION: Secreted and Trans
  / TITLE OF INVENTION: Acids Encoding t
  / FILE REFERENCE: 10466-14
  / CURRENT APPLICATION NUMBER: US/09/902
  / CURRENT FILING DATE: 2001-07-10

```

```

; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-853-376

Alignment Scores:
Pred. No.:      4.42e-36          Length:      997
Score:           398.50          Matches:     84
Percent Similarity: 57.28%       Conservative: 34
Best Local Similarity: 40.78%    Mismatches:   81
Query Match:      34.21%         Indels:        7
DB:               9              Gaps:          5

US-09-766-511B-53 (1-209) x US-09-902-853-376 (1-997)

QY      3  GluGluGlnGlnProGInSerThrGlyLysArgGlyTrpLeuSer-----LeuArgLeu 20
      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      118 AATCATCTGAACACACAATGCACAG--AGAGGATGCTTCTCTCCCAAAATGTCCTTA 174

QY      21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerIalaCysPheIleValSerCysVal 40
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      175 TGGACTGTGCTGGGATGCCCATCTATTCTCAGTGCCCTGTTTCATCACAGATGTGTT 234

QY      41 ValPhrTrpHisPheThrTrpGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      235 GTGACATTCGCATC---TTTCAAACCTGTGATGAGAAAAAAGTTTCAGCTACCTGAGAAT 291

QY      61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
      ::::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      292 TTCACAGAGCTCTCGTCTATAAATATGATGATCAGGT---TCAGTCAAGAAATTGTTGTCCA 348

QY      81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      349 TTGNACTGGGAATATTTCATACAGCTGCTACTCTTTTCTACTGACACCATTTCCTGG 408

QY      101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```
Db 409 GCGTTAGTTTAAAGAACTGCTCAGCCATGGGGCTCACCTGGTGGTTATCAACTCACAG 468
Qy 121 AlaGluGlnAsnPhelIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 469 GAGGACGAGGAATCCCTTCTCAAGAAGAACCTAAATATGAGAGATTTTTTATTGGACTG 528
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 529 TCAGACCAGGTGTCGAGGGTCAGTGGCAATGGGTGACGGCACACCTTTTGACAAAGTCT 588
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178
Db 589 CTGAGCTTCTGGGATGAGGGAGGCCCAACAACATAGCTACCTCGGAGGACTGTGCCACC 648
Qy 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
Db 649 ATGAGAGACTCTTCAACCCCAAGGCAAAATGGGAATGATGTAACCTGTTTCTCAATTAT 708
Qy 199 AsnSerIleCysGluMet 204
Db 709 TTTCGGATTTCGAAATG 726

RESULT 14
US-09-766-511b-53 (1-209) x US-09-907-824-376 (1-997)
; Sequence 376, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
```

```
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-824-376

Alignment Scores:
Pred. No.: 4.42e-36 Length: 997
Score: 398.50 Matches: 84
Percent Similarity: 57.28% Conservative: 34
Best Local Similarity: 40.78% Mismatches: 81
Query Match: 34.21% Indels: 7
DB: Gaps: 5

US-09-766-511b-53 (1-209) x US-09-907-824-376 (1-997)
```

```
Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20
Db 118 AATCATCTGAACACATGCACAGAG--AGAGATGCTTCTTCCCAAATGTTCTTA 174
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40
Db 175 TGGACTGTTGCTGGGATCCCATCTATTTCACAGTGCCTGTTTCATCACCAGATGTGT 234
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60
Db 235 GTGACATTTCCGATC---TTTCAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAAT 291
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
Db 292 TTCACAGAGCTCTCCCTGCTACAAATATGATCAGGT---TCAGTCAAGAAATTTGTGTCCA 348
Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100
Db 349 TTCAACTGGGAATATTTTCAATCCACCTGCTACTTCTTTTCTACTCACACCATTTCCCTGG 408
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
Db 409 GCGTTAAGTTTAAAGAACTGCTCAGCCATGGGGCTCACCTGGTGTATCAACTCACAG 468
Qy 121 AlaGluGlnAsnPhelIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 469 GAGGACGAGGAATTCCTTCTCAAGAAGAACCTTAAATGAGAGATTTTTTATTGGACTG 528
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 529 TCAGACCAGGTGTCGAGGGTCAGTGGCAATGGGTGACGGCACACCTTTTGACAAAGTCT 588
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178
Db 589 CTGAGCTTCTGGGATGAGGGAGGCCCAACAACATAGCTACCTCGGAGGACTGTGCCACC 648
```

QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198  
Db 649 ATGAGAGACTCTTCAACCCAGGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTAT 708  
QY 199 AsnSerIleCysGluMet 204  
Db 709 TTTTCGGATTGTGAAATG 726

RESULT 15

US-09-907-841-376

; Sequence 376, Application US/09907841

; Publication No. US20020198366A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: US/09/907,841

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 376

; LENGTH: 997

; TYPE: DNA

; ORGANISM: Homo Sapien

US-09-907-841-376

Alignment Scores: 4.42e-36 Length: 997

Pred. No.:

Score: 398.50 Matches: 84  
Percent Similarity: 57.28% Conservatives: 34  
Best Local Similarity: 40.78% Mismatches: 81  
Query Match: 34.21% Indels: 7  
DB: 9 Gaps: 5

US-09-766-511b-53 (1-209) x US-09-907-841-376 (1-997)

QY 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20  
Db 118 AAATCATCTGAAACACAATGCACAGAG--AGAGGATGCTTCTTCCCAAAATGTTCTTA 174  
QY 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
Db 175 TGGACTGTTGCTGGATCCCATCTTATTTCTCAGTGCCTGTTTTCATCACCAGATGTT 234  
QY 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 235 GTGACATTTCCGATC---TTTCAACCTGTGATGAGAAAAAGTTTCAGCTACTGAGAT 291  
QY 61 HisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 292 TTCACAGAGCTCTCTGCTACAATTTATGGATCAGGT---TCAGTCAAGAATTTTGTCCA 348  
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
Db 349 TTGAACCTGGGAATATTTTCAATCCAGCTGCTACTTTCTTTCTACTGACACCATTTCTCG 408  
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGlu 120  
Db 409 GCCTTAAGTTTAAAGAACTGCTCAGCCATGGGGCTCAGCTGGTGTATATCAACTCACAG 468  
QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
Db 469 GAGGAGCAGGAATTCCTTCCACAAAGAACTTAAATGAGAGAGTTTTATTGGACTG 528  
QY 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
Db 529 TCAGACCAGGTTGTCGAGGGTCAGTGGCAATGGTGGAGGCACACCTTTTGACAAAGTCT 588  
QY 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178  
Db 589 CTGAGCTTCTGGGATGTAGGGAGGCCCAACACATAGCTACCCCTGGAGGACTGTGCCACC 648  
QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198  
Db 649 ATCAGAGACTCTTCAACCCAGGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTAT 708  
QY 199 AsnSerIleCysGluMet 204  
Db 709 TTTTCGGATTGTGAAATG 726

Search completed: February 20, 2003, 20:41:58

Job time : 62 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 19:24:22 : Search time 1689 Seconds  
(without alignments)  
2004.060 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPOSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094339\_29300/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1.1.899\_@runat\_19022003\_094339\_29300 -NCPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_esthc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_est3.\*  
12: gb\_est4.\*  
13: gb\_est5.\*  
14: gb\_est6.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pin.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	694.5	59.6	971	13	BI696642	BI696642 603348532
2	682	58.5	504	12	BF603486	BF603486 269014 MA
3	666	57.2	470	12	BF774158	BF774158 283735 MA
4	475.5	40.8	631	11	AK007794	AK007794 Mus muscu
5	439.5	37.7	582	9	AI183967	AI183967 qd69e10.x
6	427.5	36.7	568	10	BE200183	BE200183 uq60b08.x
7	419	36.0	514	12	BF152383	BF152383 u232f08.y
8	419	36.0	558	9	AI391056	AI391056 mcl3h05.y
9	414.5	35.6	812	11	AK020363	AK020363 Mus muscu
10	400.5	34.4	403	10	AW822082	AW822082 up67g10.x
11	390.5	33.5	1200	11	AK014135	AK014135 Mus muscu
12	382	32.8	1013	9	AA868502	AA868502 ak43b01.s
13	377.5	32.4	753	9	AI949910	AI949910 wq04a07.x
14	375.5	32.2	561	10	BE650539	BE650539 UT-M-BH2
15	373.5	32.1	1032	10	BE311916	BE311916 601143328
16	363	31.2	715	10	AV716802	AV716802 AV716802
17	362	31.1	889	12	BG164873	BG164873 602343809
18	359	30.8	676	13	BI909298	BI909298 603067260
19	352	30.2	831	12	BG065735	BG065735 H3034F05-
20	350.5	30.1	582	10	AV716725	AV716725 AV716725
21	349.5	30.0	631	12	BG080418	BG080418 H3053D08-
22	348.5	29.9	498	10	BE689245	BE689245 uw51d04.y
23	342.5	29.4	660	10	BE292395	BE292395 601057892
24	342	29.4	1007	14	BQ1715169	BQ1715169 AGENCOURT
25	333.5	28.6	854	13	BI910606	BI910606 603069821
26	316.5	27.2	667	10	BE688920	BE688920 uw51d04.x
27	315	27.0	576	12	BG079007	BG079007 H3034F05-
28	314.5	27.0	334	10	BE690187	BE690187 uw65c04.y
29	314.5	27.0	738	14	BQ004812	BQ004812 UI-H-E10-
30	299.5	25.7	448	9	AA446401	AA446401 zw58d11.r
31	290.5	24.9	666	10	AW242656	AW242656 xm89a06.x
32	288.5	24.8	681	12	BF020383	BF020383 uw65c04.x
33	278.5	23.9	678	12	BF227688	BF227688 u232f08.x
34	267.5	23.0	559	10	BE120531	BE120531 UI-R-CA0-
35	267	22.9	432	10	BB863598	BB863598 BB863598
36	261	22.4	604	10	AW743449	AW743449 up67g10.y
37	259	22.2	645	13	BI328284	BI328284 602985590
38	259	22.2	714	12	BG533278	BG533278 602073733
39	254.5	21.8	366	13	BI027634	BI027634 CW3-WT029
40	251.5	21.6	458	10	BE632986	BE632986 uv71e02.y
41	251.5	21.6	816	14	BQ018186	BQ018186 UI-H-DT1-
42	249	21.4	668	14	BQ674807	BQ674807 AGENCOURT
43	249	21.4	785	14	BQ648893	BQ648893 AGENCOURT
44	249	21.4	886	14	BQ651695	BQ651695 AGENCOURT
45	249	21.4	914	9	AL564408	AL564408 AL564408

ALIGNMENTS

RESULT 1  
BI696642  
LOCUS 603348532F1 NC1\_CGAP\_Mam2 Mus musculus cDNA clone IMAGE:5376247 5',  
DEFINITION mRNA sequence.  
ACCESSION BI696642  
VERSION BI696642.1 GI:15659271  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 971)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)





```
Db 7 GCTGAAGTGCACACACACACACATTAAATCTAACCTGCTTACAGTGAAGGACAAAGGTCACA 66
Qy 75 Ala-----TTPGLYCYSProAlaSerTrpLysSerPheGlySerCysTyrPh 92
Db 67 NGAAGAAGATTGGGGATGTTGCCACAGTACCTGCAAGCGCTTGGTCCAGCTGCTACTT 126
Qy 92 eLleSerSerGluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAl 112
Db 127 TATTCTTCTGAAGAAGATTCTGGGCTAAGAGTGAAGAGTGCATGGGATGGGAGC 186
Qy 112 aHisLeuValValPheAsnThrGluAlaGluGlnAsnPheLeValGlnGlnLeuAsnG 132
Db 187 TCACCTGGTGTGATCACACACAGAAACAGACAGAGGATTCATTATCCAGCAGCTGAATA 246
Qy 132 uSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIl 152
Db 247 AACATTTCTTATTCTTGGGACTCTCAGACCCCAAGGAATGGCAACTGGCAATGGAT 306
Qy 152 eAspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSe 172
Db 307 TGATCAGACACCTTACAAGGAAATGTGAGATTTTGGCCCAAAATGAACCCAACTTTC 366
Qy 172 rAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVa 192
Db 367 TCACAGGAATGCTGCTCAGTTGTTTCTGGGATGGGAGAGGATGGGGCTGGAATGATGT 426
Qy 192 lLeCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 427 TTTCTGTGATCTTAAAGGAATCAATATGTGAGATGAAGAAGATTTACCTA 478

RESULT 3
BF774158 470 bp mRNA linear EST 25-APR-2001
LOCUS 283735 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF774158
ACCESSION BF774158.1 GI:12122058
VERSION EST.
KEYWORDS cow.
SOURCE Bos taurus
ORGANISM Bos taurus
REFERENCE Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
AUTHORS Casals,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.B., Heaton,M.P., Laegreid,W., Kohrer,G.A., Chitko-McKown,C.G.,
Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACACCTATGACCAT
BACKWARD: GTTTCACGTCACGACG
Plate: 80 row: M column: 16
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
Location/Qualifiers
1..470
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
```

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinosus muscle, and fetal  
longissimus muscle."

BASE COUNT 138 a 84 c 122 g 124 t 2 others  
ORIGIN

## Alignment Scores:

Pred. No.: 2e-66 Length: 470  
Score: 686.00 Matches: 116  
Percent Similarity: 85.99% Conservative: 19  
Best Local Similarity: 73.89% Mismatches: 18  
Query Match: 57.17% Indels: 4  
DB: 12 Gaps: 1

US-09-766-511B-53 (1-209) x BF774158 (1-470)

```
Qy 52 LysArgLeuSer-GluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyTh 71
Db 1 AAGAGGCTGTCNTGAAGTGCACACACACCATTTCAAATTAACCTGCTTCAGTGAAGGAC 60
Qy 71 rLysValProAla-----TTPGLYCYSProAlaSerTrpLysSerPheGlySerS 89
Db 61 AAGGGTGACANGAAAGATTGGGGATGTTGCCAGGTACTCTGAAGCCGCTTTGGTTCCA 120
Qy 89 erCysTyrPheIleSerSerGluGluLysValTrpSerLysSerGluGlnAsnCysVal 109
Db 121 GCTGCTACTTATTCTCTGAAGAGAAATTTCTGGCTAAGACTGAGCAGAACTGCATTG 180
Qy 109 LuMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGln 129
Db 181 GGATGGGAGCTCACTTGGTGTATCAACACAGAAACAGACAGAGGATTTTCATTATCCAGC 240
Qy 129 InLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTr 149
Db 241 AGCTGAATAAAACATTTTCTTATTTTCTGGGACTCTCAGACCCCAAGGAATGGCAACT 300
Qy 149 rpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGlu 169
Db 301 GGCATGGATTGATCAGACACCTTACAAGGAAATGTGCAGATTTTGGCACCACAAATGAAC 360
Qy 169 roAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGly 189
Db 361 CCAACTTTTCTGCAGAGAAATGTCTCAGTTGTTTCTGGGATGGGAGGAGGGGGCT 420
Qy 189 rpAsnAspValIleCysGlnThrArgArgAsnSerIleCysGluMet 204
Db 421 GGAATGATGTTTCTGTGATCTTAAAGGAATCAATATGTGAGATG 467
```

## RESULT 4

AK007794 631 bp mRNA linear HTC 19-JAN-2002  
LOCUS AK007794  
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length  
enriched library, clone:1810046124, similar to DENDRITIC CELL  
IMMUNORECEPTOR, full insert sequence.

## ACCESSION

AK007794.1 GI:12841567

## VERSION

HTC; CAP trapper.

## KEYWORDS

Mus musculus (Strain:C57BL/6J) 10 day old male pancreas cDNA to

## SOURCE

mRNA, clone\_lib:RIKEN full-length enriched mouse cDNA library

clone:1810046124.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

PUBMED 10349636

## REFERENCE

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20493374  
11042159

## AUTHORS

3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsudaira,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

4  
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., De Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gusincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohtsuki,S. and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

## AUTHORS

5 (bases 1 to 631)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sugabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9226, Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAAAGCTCTTTTTTTTTTNN 3'], cDNA was

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

US-09-766-511b-53 (1-209) x AK007794 (1-631)  
Qy 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20  
Db 94 ATGCTTCAGGAAGACAGCACAAGGG---AAGCAGCTCTCTGG---TCCCTGAGACTC 147  
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
Db 148 TGGTCAGCTGCTGTGATTTCTATCTACTCTCAGCACCTGTTTCATTGCGAGCTGTGTG 207  
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 207 -----207  
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 208 -----GACAAAGTC-----TGGAGCTGTGCCCCA 231  
Qy 81 AlasSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer-----GluGluLys 98  
Db 232 AAGGATTGGGAAGTATTGTGTTCCCACTGCTACTTGGTCCACACAGCTTTTTCATCAGCA 291  
Qy 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118  
Db 292 TCITGGAAACAGAGTGGAGAGAACTGCTCCCGCATGGGTGCTCATCTAGTGTGTATCCAT 351  
Qy 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeu 138  
Db 352 AGCCAGGAGGAGGAGGAGGATTTTCATCACTGGGATCTTGGACATTCATGCTCTATTATTA 411  
Qy 139 GlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGlu 158

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

US-09-766-511b-53 (1-209) x AK007794 (1-631)  
Qy 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20  
Db 94 ATGCTTCAGGAAGACAGCACAAGGG---AAGCAGCTCTCTGG---TCCCTGAGACTC 147  
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
Db 148 TGGTCAGCTGCTGTGATTTCTATCTACTCTCAGCACCTGTTTCATTGCGAGCTGTGTG 207  
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 207 -----207  
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 208 -----GACAAAGTC-----TGGAGCTGTGCCCCA 231  
Qy 81 AlasSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer-----GluGluLys 98  
Db 232 AAGGATTGGGAAGTATTGTGTTCCCACTGCTACTTGGTCCACACAGCTTTTTCATCAGCA 291  
Qy 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118  
Db 292 TCITGGAAACAGAGTGGAGAGAACTGCTCCCGCATGGGTGCTCATCTAGTGTGTATCCAT 351  
Qy 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeu 138  
Db 352 AGCCAGGAGGAGGAGGAGGATTTTCATCACTGGGATCTTGGACATTCATGCTCTATTATTA 411  
Qy 139 GlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGlu 158

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAAAGCTCTTTTTTTTTTNN 3'], cDNA was

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGCGCGCAATTAAATTCGAGTTAAATTAATCCGCCCCC 3'], cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

## FEATURES

Location/Qualifiers  
Source  
1..631

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:1810046124"  
/db\_xref="MGD:MGI:1901987"  
/db\_xref="taxon:10090"  
/clone="1810046124"  
/sex="male"  
/tissue\_type="pancreas"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 day old"  
94..624

## CDS

/note="data source:SPTR, source key:Q9QZL5, evidence:ISS putative similar to DENDRITIC CELL IMMUNORECEPTOR"

/codon\_start=1  
/protein\_id="BAB25260.1"  
/db\_xref="GI:12841568"  
/db\_xref="MGD:MGI:1917060"  
/translation="MVGRQLQKAVSWLSRLMSAAVISILLSTCTFIASCVDKVMSC  
CPDKMLFGSHCYLPTVFSSANKSENCRRMGAHLVYIHSSEEDDITGLDIHA  
AYFTGLMDTGHQWQWDDTPYEYTFWHNGEPSSDNRKCVTVYVRRNIGWNDIS  
CNLKQKSVCCOMKINL"

BASE COUNT 174 a 129 c 157 g 171 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.76e-44 Length: 631  
Score: 475.50 Matches: 92  
Percent similarity: 57.82% Conservative: 30  
Best Local Similarity: 43.60% Mismatches: 52  
Query Match: 40.82% Indels: 37  
DB: 11 Gaps: 6

```
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 GGGTGTGGGAT---ACAGGCCATCGCAATGGCAATGGTTGATCAGACACCATATGAA 468

QY 159 LysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSer 178
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 469 GAAAGTGTCACTCTGGCACAATGGTGAGCCCGCAGCGAGTGCACAAATGAAAATGTGTACA 528
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 529 GTATATTACCGTCGGAATATTGGATGGGCGTGAATGATATCTCTTGCAATCTTAAACAG 588
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 199 AsnSerIleCysGluMetAsnLysIleTyrLeu 209
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 589 AAGTCAGTTTGTGAGATGAAGAAAATAAACTTA 621

RESULT 5
AL183967/c
LOCUS
DEFINITION
  qd9e10.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1734762
  3' similar to SW:LECH_MOUSE P34927 ASIALOGLYCOPROTEIN RECEPTOR 1 ;,
  mRNA sequence.
ACCESSION
  AL183967
VERSION
  AL183967.1 GI:3734605
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Glibco
High quality sequence stop: 453.
FEATURES
  source
    1..582
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:1734762"
      /clone_lib="Soares_testis_NHT"
      /sex="male"
      /lab_host="DH10B"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
      was prepared from mRNA obtained from Clontech Laboratories
      , Inc., and primed with a Not I - oligo(dT) primer [5',
      TGTTACCAATCTGAAGTGGGAGCGGCCCAATTTTTTTTTTTT 3'].
      Double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not I
      and Eco RI sites of the modified pT7T3 vector. Library
      went through one round of normalization to Cot5, and was
      constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 175 a 131 c 105 g 171 t
ORIGIN
```

Alignment Scores:  
Pred. No.: 3,41e-40 Length: 582  
Score: 439.50 Matches: 78  
Percent Similarity: 68.57% Conservative: 18  
Best Local Similarity: 55.71% Mismatches: 38  
Query Match: 37.73% Indels: 6  
DB: Gaps: 1

```
US-09-766-511b-53 (1-209) x AI183967 (1-582)

QY 76 TrpGlyCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95
      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 579 TGGAGCTGCTGCCCAACCCCTTGGACTTCATTTCAGTCTAGTCTAGTATTATTCTACT 520
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 519 GGGATGCAATCTTGAGCTAAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG 460
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 116 ValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSer 135
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 459 GTGATCAACACACGAGGGAAGACAGGATTTCATTCATTCAGAAATCTGAAAAGAAATTTCTTCT 400
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 399 TATTTTCTGGGCTGTGAGATCCAGGGGTCCGCCACATTTGGCAATGGTTGGACCGACA 340
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 156 ProTyrGluLysAsnVal-----ArgPheTrpHisLeuGlyGluProAs 170
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 339 CCATCAATGAAAAATGTCACGTGAGTATAGAATGAGATTCTGGCAGCTCAGGTGAACCCAA 280
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 170 nHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAs 190
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 279 TAACCTTGATGAGCGTTGTGCGATAATAAATTTCCGTTCTTCAGAGAAGATGGGCTGGAA 220
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
QY 190 nAspValIleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
      : : : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 219 TGACATTACTGTCATGTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 162

RESULT 6
BE200183/c
LOCUS
DEFINITION
  ug60b08.x1 Soares_mammary_gland_NMLKG Mus musculus cDNA clone
  IMAGE:1546743 3' similar to TR:Q9Z2H6 Q9Z2H6 C-TYPE LECTIN. ;, mRNA
  sequence.
ACCESSION
  BE200183
VERSION
  BE200183.1 GI:8712352
KEYWORDS
  EST.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 568)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: ug60b08.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMGI Consortium (info@image.llnl.gov) for further information.
MGI:952091
High quality sequence stop: 464.
FEATURES
  source
    1..568
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /clone="IMAGE:1546743"
      /clone_lib="Soares_mammary_gland_NMLKG"
      /sex="female (lactating)"
      /tissue_type="mammary gland"
      /lab_host="DH10B"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
      polylinker; 1st strand cDNA was prepared from mammary
      gland tissue from a lactating female, and was then primed
      with a Not I - oligo(dT) primer. Double-stranded cDNA was
      ligated to Eco RI adaptors (Pharmacia), digested with Not
      I and cloned into the Not I and Eco RI sites of the
      modified pT7T3 vector. Library is normalized. Library
      was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 155 a 128 c 110 g 175 t
```

ORIGIN

Alignment Scores:  
Pred. No.: 7,92e-39 Length: 568  
Score: 427.50 Matches: 78  
Percent Similarity: 62.90% Conservative: 39  
Best Local Similarity: 41.94% Mismatches: 60  
Query Match: 36.70% Indels: 9  
DB: 10 Gaps: 7

US-09-766-511B-53 (1-209) x BE200183 (1-568)

QY 30 LeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPheThrTyrGlyGlu 49  
DB 567 CTCTGGCAATCACAATCTTAGTCTTTATCAT--TATTTTCAAAAGTACTCTCAA 511  
QY 50 ---ThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThrCysPheSer 68  
DB 510 CTTCTTGAAGAAATAAAGCTCCAAAATAATATGTCACATGAATTCAGCTGCACAAA 451  
QY 69 GluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrpLysSerPhe 86  
DB 450 AGTGTTTCCACCATGAAGACAAAGCTCTGGAGCTGTTGCCCAAGGATGGAGGCTATT 391  
QY 87 GlySerSerCysTyrPheIleSer-----SerGluGluLysValTrpSerLysSerGlu 104  
DB 390 GGTGCCACTGCTACTTGGTTCCTCCACAGTCTTCTCATCAGCATCTTGGAAAGAGTGAG 331  
QY 105 GluAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGluAsn 124  
DB 330 GAGACTGCTCCCGATGGTGCTCATCTAGTGGTGATCCAAAGCCAGAGAGCAGGAT 271  
QY 125 PheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGln 144  
DB 270 TTCATCACTGGGATCTGGACACTCATGCTGCTTATTTATAGGCTTGGGAT---ACA 214  
QY 145 GlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPheTrp 164  
DB 213 GGCCATGGCAATGGCAATGGTGTTCAGACACCATATGAAGAAGTATCACATTTCTGG 154  
QY 165 HisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPhe---TrpLys 183  
DB 153 CACAATGGTGAGCCAGCAGTGGCAATGAAATAATGTCTACAATAATTTACCGTTGGAG 94  
QY 184 ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCysGlu 203  
DB 93 ---ACTGGATGGGCTGGAAACATATCTCTGCACATCTTAAACAGAGTCACTTGTCAG 37  
QY 204 MetAsnLysIleTyrLeu 209  
DB 36 ATGAAGAAATAAATCTTA 19

RESULT 7  
BF152383 514 bp mRNA linear EST 29-DEC-2000  
LOCUS uz32f08.y1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3670791 5',  
similar to TR:Q9QZ15 Q9QZ15 DENDRITIC CELL IMMUNORECEPTOR. ;, mRNA  
sequence.

ACCESSION BF152383

VERSION BF152383.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

COMMENT Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

FEATURES

source  
1. .514  
Location/Qualifiers  
Seq primer: -40RP from Gibco  
High quality sequence stop: 436.

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3670791"  
/clone\_lib="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 161 a 96 c 117 g 140 t

ORIGIN

Alignment Scores:  
Pred. No.: 6,45e-38 Length: 514  
Score: 419.00 Matches: 73  
Percent Similarity: 64.97% Conservative: 29  
Best Local Similarity: 46.50% Mismatches: 53  
Query Match: 35.97% Indels: 2  
DB: 12 Gaps: 2

US-09-766-511B-53 (1-209) x BF152383 (1-514)

QY 54 LeuSerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal 73  
DB 47 ATAAAGAACTGAACATATACTACTGAATTGGAGTGTACAAAATGGCTTCACTCTTGAAGAC 106  
QY 74 ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIle 93  
DB 107 AAAGTCTGGAGCTGTGCCCAAGGATTTGAAGCCGTTTGGTTCCTACTGCTACTTCACT 166  
QY 94 SerSerGlu---GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAla 112  
DB 167 TCAACTGACTTGGTGGCATCTTGGAAATGAGAGTAAGAGAACTGCTTCCACATGGGTGCT 226  
QY 113 HisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGlu 132  
DB 227 CATCTGGTGTGATCCACAGCCAGGAGAACAGAGATTTCATCACTGGGATCCTGGACACT 286  
QY 133 SerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIle 152  
DB 287 GGACTGCTTATTTATAGACTTTCAAAATCCA---GGTGATCAACAATGGCAATGGATT 343  
QY 153 AspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSer 172  
DB 344 GATCAGACCCGTCAGCATGATAACCATCCACATCTGCGACAAAGGTGAGCTAGCAGTGAC 403  
QY 173 AlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVal 192  
DB 404 AATGAACAGTGTGTTTAAATAATCATCTGCAGAGTACTGGATGGGCTGGAGTGATATC 463  
QY 193 IleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
DB 464 CCTTGCAGTGATAACAGAACTCAATTTGCCATGTGAAAAAATAATATCTTA 514

RESULT 8

AI391056

LOCUS

AI391056 558 bp mRNA linear EST 15-MAR-2000

**DEFINITION**  
 mcl3h05.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone  
 IMAGE:348441 5' similar to SW:MANR\_HUMAN P22897 MACROPHAGE MANNOS  
 RECEPTOR PRECURSOR. ; mRNA sequence.

**ACCESSION**  
 AI391056

**VERSION**  
 AI391056.1 GI:4217063

**KEYWORDS**  
 EST.

**SOURCE**  
 house mouse

**ORGANISM**  
 Mus musculus

**REFERENCE**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 538)

**AUTHORS**  
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person  
 , B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter  
 , E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.

**TITLE**  
 The WashU-NCI Mouse EST Project 1999

**JOURNAL**  
 Unpublished (1999)

**COMMENT**  
 Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 This read is a RESEQUENCE of a previously sequenced mouse clone  
 This read has been verified (found to hit its original self in the  
 correct orientation)  
 MGI:220241  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 491  
 POLYA=No.

#### FEATURES

Source  
 1. .558  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:348441"  
 /clone\_lib="Soares mouse p3NMF19.5"  
 /dev\_stage="19.5 dpc total fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Vector: pT73D (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5',  
 TCTTACCAATCGAAGTGGAGCGCGCATTTTGTGTTTGT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA was kindly provided by  
 Dr. Minoru Ko (Wayne State University)."  
 BASE COUNT 171 a 108 c 129 g 150 t

#### Alignment Scores:

Pred. No.: 7,33e-38 Length: 558  
 Score: 419.00 Matches: 73  
 Percent Similarity: 64.97% Conservative: 29  
 Best local Similarity: 46.50% Mismatches: 53  
 Query Match: 35.97% Indels: 2  
 DB: 9 Gaps: 2

US-09-766-511B-53 (1-209) x AI391056 (1-558)

Qy 54 LeuSerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal 73  
 ::: |||||:::  
 Db 14 ATAAAGAACTGAATATACTAATGGAGTGATACAAAATGGCTTCACCTCTTGAAGAC 73  
 Qy 74 ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIle 93  
 ||| |||||:::  
 Db 74 AAAGTCGGAGCTGTTGCCCAAGAGATTGGAAGCCGCTTTGGTTCTTCTACTGCTCACTTCACT 133

Qy 94 SerSerGlu---GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAla 112  
 |||||:::  
 Db 134 TCAACTGACTTGGTGGCATCTTGAATCAGAGTAAGGAGAAGCTGCTCCACATGGTGCT 193  
 Qy 113 HisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGlu 132  
 |||||:::  
 Db 194 CATCTGGTGGTGCATCCACAGCCAGGACAGGAGATTTCATCACTGGGACTCTGGACACT 253  
 Qy 133 SerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIle 152  
 |||||:::  
 Db 254 GGTACTGCTTAATTTATAGGACTTTCAAATCCA---GGTATCAACAATGGCAATGGATT 310  
 Qy 153 AspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSer 172  
 |||||:::  
 Db 311 GATCAGACACCGTACGATGATAATACCACATCTCTGGCACAAAGCTGAGCCCTAGCAGTGAC 370  
 Qy 173 AlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVal 192  
 |||||:::  
 Db 371 AATGAACAGTGTGTATATAAATCATCGTCAGAGTACTGGATGGGCTGGAGTGATATC 430  
 Qy 193 IleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 |||||:::  
 Db 431 CTTGCACTGATAACAGAACTCAATTGCCATGTGAAAAAATATACTTA 481

RESULT 9  
 AK020363 812 bp mRNA linear HTC 19-JAN-2002

LOCUS  
 AK020363 Mus musculus adult male epididymis cDNA, RIKEN full-length enriched  
 library, clone:9230119016:C-type (calcium dependent, carbohydrate  
 recognition domain) lectin, superfamily member 6, full insert  
 sequence.

ACCESSION  
 AK020363

VERSION  
 AK020363.1 GI:12860929

KEYWORDS  
 HTC; CAP trapper.

SOURCE  
 Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,  
 clone\_lib:RIKEN full-length enriched mouse cDNA library  
 clone:9230119016.

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
 1 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 PUBMED 10349636

REFERENCE  
 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 PUBMED 11042159

REFERENCE  
 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multipillar sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 PUBMED 11076861

REFERENCE  
 4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamakawa,I.,  
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bernaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamli, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

5 (bases 1 to 812)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-72 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGACTCTTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES  
source  
1. .812  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:9230119016"  
/db\_xref="MGI:1911630"  
/db\_xref="taxon:10090"  
/clone="9230119016"  
/sex="male"  
/tissue\_type="epididymis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1. .812  
/gene="Clec3f6"  
misc\_feature  
1. .812

/gene="Clec3f6"  
/note="C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 6  
data source:MGI:1349412, evidence:ISS"  
/db\_xref="MGI:1349412"  
BASE COUNT 227 a 160 c 176 g 249 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4, 35e-37 Length: 812  
Score: 414.50 Matches: 74  
Percent Similarity: 64.74% Conservative: 27  
Best Local Similarity: 47.44% Mismatches: 48  
Query Match: 35.58% Indels: 7  
DB: 11 Gaps: 5  
US-09-766-511b-53 (1-209) x AK020363 (1-812)  
QY 57 LeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrrp 76  
Db 2 TTAGAGTCAATTGAGATGAGACTTTGGGTATCTTGGCACCTGAGCACAAGTC-----TGG 55  
QY 77 GlyCysCysProAlaSerTrrpLysSerPheGlySerSerCysTyrPheIleSer----- 94  
Db 56 AGCTGTTGGCCCAAGGATTGGAGGCTATTGTTCCCACTGTCTTCCACAGTT 115  
QY 95 SerGluGluLysValTrrpSerLysSerGluGluInAsnCysValGluMetGlyAlaHisLeu 114  
Db 116 TCITTCATCACATCTTGGACACAGAGTGGAGAGAACTGTCCTCCGCATGGGTGCTCATCTA 175  
QY 115 ValValPheAsnThrGluAlaGluGluInAsnPheIleValGluGluInLeuAsnGluSerPhe 134  
Db 176 GTGCTGATCCAAAGCCAGGAGAGGAGGATTTCATCTCTGGGATCTTGGACATCATGCT 235  
QY 135 SerTyrPheLeuGlyLeuSerSerProGluGlyAsnAsnAsnTrpGlnTrpIleAspLys 154  
Db 236 GCTATTATTATAGGGTTGTGGGAT---ACAGCCCAATGGCAATGGGTGTATCATG 292  
QY 155 ThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGlu 174  
Db 293 ACACCATATGAAGAACTATCATCTTGGCACAAATGGTGAGCCAGCAGTGCATGAA 352  
QY 175 GlnCysAlaSerIleValPhe---TrpLysProThrGlyTrpGlyTrpAsnAspValIle 193  
Db 353 AAATGCTGCTACAATAATTTACCGTTGGAAG---ACTGGATGGGCTGGAACGATATCTCT 409  
QY 194 CysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
Db 410 TGCAGTCTTAACAGAGAGTCAGTTTGTCTAGATGAAGAAATAAATTA 457

RESULT 10  
AW822082/c  
LOCUS  
DEFINITION  
up67g10.x1 Soares\_mouse\_NMIE Mus musculus cDNA clone IMAGE:2781186  
3' similar to TR:Q922H6 Q922H6 C-TYPE LECTIN. ;, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Mus musculus  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 403)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Other\_ESTs: up67g10.y1  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
This clone is available royally-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:1037606









Percent Similarity:	57.92%	Conservative:	26
Best Local Similarity:	43.72%	Mismatches:	67
Query Match:	32.79%	Indels:	12
DB:	9	Gaps:	2

US-09-766-511B-53 (1-209) x AA868502 (1-1013)

Qy 29 AlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPheThrTyrGly 48  
:::lll ::::lll ::::lll ::::lll ::::lll ::::lll ::::lll ::::lll ::::lll ::::lll  
Db 679 AGCTCAGGACCGAGTGCCTCACAAATTTATGTATG----- 644

Oy 49 GluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeu-----ThrCys 66  
Db 643 CAAACGGTCAGAGTGCTCGCAAGTTACGGAGTTCAACGGTCTCACTCAAGCCGACTGC 584

Qy 67 PheSerGluGlyThrLysValProAlaTrpGlyCysCysProAlaSerTrpLysSerPhe 86  
::: ||| ||||| ||| |||||  
Db 583 GTCTGTAAAGATAAGTACATAGAAGATTGGAGTGTCTGCCAACCCC-TGGACTTCATTC 525

Qy 87 GlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSerGluGlnAsn 106  
 |||||:::|||||::: :: |||:::|||||:::  
 Db 524 AGTCCTAGTGTCTACTATATTCTACTGG-GATGCATCTTGGACTAAGAGTCAAGACC 466

Qy 107 CysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIle 126  
Db 465 TCCTCTGTGATGTGAGCTGATCTGGTGGTGATCACCCAGGGAAGACAGGATTTCATC 406

	ValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsn	146
Qy	127 ValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsn	146
	:::::     :::	
Db	405 ATTCAGAACTCGAAGAATAATCTTCTATTATTTCTGGGCTGTCCAGATCCAGGGGTCGG	346

[illegible]

**Qy**    I67 GlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGly  
         ||||| :||||| :: |::: ||||| ||| |||  
**Dd**    285 GTGAACCCCAATACTTGTATGAGCGTTGGTGATAATAAATTTC CGTTCTTCAGACGAA 2266

QY 18/ trpgyirtpasnaspvalliicysgciuurnrargasnrlliecysgiuimetasnlys 209  
||||| ||||| ::||| ::||| ::||| ::||| ::||| |||||  
Db 225 TGGGCTGGAATGACATTCACTGTGCATGTACCTCAGAAGTCAATTGCCAAGTGAAGAAG 166

QY 207 I I E I Y I L E D 203  
 |||||:::  
 Db 165 ATCTACATC 157

RESULT 13	
A1949910/c	
LOCUS	753 bp mRNA linear EST 08-MAR-2000
DEFINITION	wq04a07.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2470260 3' similar to SW:LECI_MOUSE P24721 ASIALOGLYCOPROTEIN RECEPTOR 2 ; , mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 753)  
 NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 1065 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 449.

	FEATURES	SOURCE
1	100% Pure	100%
2	100% Pure	100%
3	100% Pure	100%
4	100% Pure	100%
5	100% Pure	100%
6	100% Pure	100%
7	100% Pure	100%
8	100% Pure	100%
9	100% Pure	100%
10	100% Pure	100%
11	100% Pure	100%
12	100% Pure	100%
13	100% Pure	100%
14	100% Pure	100%
15	100% Pure	100%
16	100% Pure	100%
17	100% Pure	100%
18	100% Pure	100%
19	100% Pure	100%
20	100% Pure	100%
21	100% Pure	100%
22	100% Pure	100%
23	100% Pure	100%
24	100% Pure	100%
25	100% Pure	100%
26	100% Pure	100%
27	100% Pure	100%
28	100% Pure	100%
29	100% Pure	100%
30	100% Pure	100%
31	100% Pure	100%
32	100% Pure	100%
33	100% Pure	100%
34	100% Pure	100%
35	100% Pure	100%
36	100% Pure	100%
37	100% Pure	100%
38	100% Pure	100%
39	100% Pure	100%
40	100% Pure	100%
41	100% Pure	100%
42	100% Pure	100%
43	100% Pure	100%
44	100% Pure	100%
45	100% Pure	100%
46	100% Pure	100%
47	100% Pure	100%
48	100% Pure	100%
49	100% Pure	100%
50	100% Pure	100%
51	100% Pure	100%
52	100% Pure	100%
53	100% Pure	100%
54	100% Pure	100%
55	100% Pure	100%
56	100% Pure	100%
57	100% Pure	100%
58	100% Pure	100%
59	100% Pure	100%
60	100% Pure	100%
61	100% Pure	100%
62	100% Pure	100%
63	100% Pure	100%
64	100% Pure	100%
65	100% Pure	100%
66	100% Pure	100%
67	100% Pure	100%
68	100% Pure	100%
69	100% Pure	100%
70	100% Pure	100%
71	100% Pure	100%
72	100% Pure	100%
73	100% Pure	100%
74	100% Pure	100%
75	100% Pure	100%
76	100% Pure	100%
77	100% Pure	100%
78	100% Pure	100%
79	100% Pure	100%
80	100% Pure	100%
81	100% Pure	100%
82	100% Pure	100%
83	100% Pure	100%
84	100% Pure	100%
85	100% Pure	100%
86	100% Pure	100%
87	100% Pure	100%
88	100% Pure	100%
89	100% Pure	100%
90	100% Pure	100%
91	100% Pure	100%
92	100% Pure	100%
93	100% Pure	100%
94	100% Pure	100%
95	100% Pure	100%
96	100% Pure	100%
97	100% Pure	100%
98	100% Pure	100%
99	100% Pure	100%
100	100% Pure	100%

```

1: 1723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2470260"
/clone.lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH108"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site.1: Not I; Site.2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIds 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Ronaldo."

```

BASE COUNT	230 a	171 c	138 g	212 t	2 others
ORIGIN					

Alignment Scores:		
Pred. No.:	7,05e-33	753
Score:	377.50	67
Percent Similarity:	45.93%	Conservative: 22
Best Local Similarity:	49.63%	Mismatches: 45
Query Match:	32.40%	Indels: 1
DB:	9	Gaps: 1

US-09-766-511B-53 (1-209) x AI949910 (1-753)

Qy 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95

Qy 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115

Db	672	GAATCAGCATCTGGGCAAGACAGTGTGCTAGATGNAAGGCTCACCTGCTG	613
Ov	116	ValpheASnThrGlnAlaGluGlnAspPheTleValGlnGlnIleAsnGlnSerPheSer	135

Db 612 GTGATAAACACTCAAGAAGACGAGGATTTCATCTTCCAGAATCTGCAAGAAGAACTGCT 553

QY 130 141 152 163 174 185 196 207 218 229 240 251 262 273 284 295 306 317 328 339 350 361 372 383 394 405 416 427 438 449 460 471 482 493 504 515 526 537 548 559 570 581 592 603 614 625 636 647 658 669 680 691 702 713 724 735 746 757 768 779 790 801 812 823 834 845 856 867 878 889 900 911 922 933 944 955 966 977 988 999  
 Db 552 TATTTTGTGGGGCTCTCAGATCCAGAGGTTCAGGCACATTGGCCATGGGTTCATCAGACA 493

[illegible]

Qy 176 CysAlaSerIleValPheTrpLys---ProThrGlyTrpGlyTrpAsnAspValIleCys 194

Db 432 TGCCTTGCTAAATTTTCGTAAATCATCCCAAAAGATGGGGCTGGAATGATGATTAATTGT  
 Qy 195 GluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 372 CTTGGTCTCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACTTA 328

[illegible]



Alignment Scores:  
Pred. No.: 3.33e-32 Length: 1032  
Score: 373.50 Matches: 78  
Percent Similarity: 57.14% Conservative: 34  
Best Local Similarity: 39.80% Mismatches: 77  
Query Match: 32.06% Indels: 7  
DB: 10 Gaps: 5

us-09-766-511b-53 (1-209) x BE311916 (1-1032)

Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20  
Db 130 AAATCATCTGAACACAAATGCACAGAG--AGAGGATGCTTCTCTCCCAAAATGTCTTA 186

Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuSerAlaCysPheIleValSerCysVal 40  
Db 187 TGGACTGTGCTGGATCCCAATCCCTATTTCTCAGTCCCTGTTTCATCACCAGATGTGT 246

Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 247 GTGACATTTCCGATC---TTTCAAAACCTGTGATCAGAAAAAGTTTCAGCTACCTGAGAAT 303

Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 304 TTCACAGAGCTCTCCTGCTACAATATATGATCAGGT---TCAGTCAAGAATTTGTGTCCA 360

Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
Db 361 TTGAACCTGGGAATATTTTCAATCCAGCTGCTACTTCTTTCTACTGACACCATTTCTCTGG 420

Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
Db 421 GCGTTAAGTTTAAAGAACTGCTCAGCCATGGGGCTCACCTGGTGTATCAACTCACAG 480

Qy 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
Db 481 GAGGAGCAGGAATTCCTTCTCAGAGAAACCTAAATGAGAGAGTTTATTATTGGACTG 540

Qy 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
Db 541 TCAGACCAGGTGTCGAGGGTCACTGGCAATGGGTGGCGGCACACCTTTGACAAAGTCT 600

Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178  
Db 601 CTGAGCTTCGGGATGTAGGGAGGCCCAACACATACCATACCTGGAGGACTGTGCCACC 660

Qy 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCys 194  
Db 661 ATGAGAGACTCTTCAAACCCCAAGGCACAAATGGACTGATGTAACCTGT 708

Search completed: February 20, 2003, 20:39:33  
Job time : 1695 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:38:30 : Search time 36 Seconds  
(without alignments)  
773.594 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMEQOQPQSTERRGWSLRL.....NDVICETRRNSICENMKIYL 209

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	ID	Description
1	1165	100.0	209	22	AAU00479	Human TANGO 405 pr
2	885	76.0	162	20	AAU27446	Human SDCMP3 polyp
3	793	68.1	209	19	AAW63010	Mouse dectin-2
4	793	68.1	209	20	AAU27447	Mouse SDCMP3 polyp
5	793	68.1	209	22	AAU00482	Murine TANGO 405 a
6	679	58.3	167	19	AAW63022	Mouse dectin-2 ext
7	659	56.6	175	19	AAW63018	Mouse dectin-2 iso
8	637	54.7	213	22	AAW63018	Amino acid sequenc
9	630	54.1	218	22	AAE12079	Dendritic cell (DC
10	617	53.0	211	22	AAE12080	Dendritic cell (DC

11	592.5	50.9	168	19	AAW63020	Mouse dectin-2 iso
12	591	50.7	145	19	AAW63017	Mouse dectin-2 iso
13	521	44.7	178	22	AAU00480	Murine TANGO 405 p
14	493	42.3	134	19	AAW63019	Mouse dectin-2 iso
15	459.5	39.4	148	23	AAU98015	Human partial dend
16	459	39.4	134	23	AAU98013	Human partial dend
17	458	39.3	237	20	AAW88124	Primate DCMP1 C-le
18	458	39.3	237	23	AAU98014	Human dendritic ce
19	458	39.3	237	23	ABW90277	Human polypeptide
20	458	39.3	243	22	AAU19814	Human novel extrac
21	458	39.3	243	23	ABP48034	Human polypeptide
22	458	39.3	246	22	AAU19653	Human novel extrac
23	458	39.3	246	23	ABP47873	Human polypeptide
24	436	37.4	238	20	AAW88128	Rodent DCMP1 C-lec
25	398.5	34.2	219	20	AAU13403	Amino acid sequenc
26	398.5	34.2	219	22	AAU29035	Human PRO polypept
27	398.5	34.2	219	22	AAU74697	Human membrane ass
28	398.5	34.2	219	22	AAW80271	Human PRO244 prote
29	390.5	33.5	214	22	AAW97225	Murine C-type lect
30	334.5	28.7	85	19	AAW69237	Mouse dectin-2 His
31	252	21.6	287	18	AAW15246	Asialoglycoprotein
32	251.5	21.6	316	20	AAW88125	Primate DCMP2 C-le
33	250	21.5	273	20	AAW88129	Variant primate DC
34	250	21.5	319	23	ABG66703	Human novel polype
35	249	21.4	229	18	AAW15251	Asialoglycoprotein
36	249	21.4	270	18	AAW15252	Asialoglycoprotein
37	249	21.4	342	23	ABW95436	Human angiogenesis
38	249	21.4	342	23	ABW84830	Human PRO7223 prot
39	249	21.4	457	20	AAW73633	Human secreted pro
40	249	21.4	457	23	ABP61864	Human polypeptide
41	249	21.4	542	20	AAW73634	Human secreted pro
42	249	21.4	542	23	ABP61865	Human polypeptide
43	249	21.4	547	22	AAW27236	Human EXMAB-14 SEQ
44	249	21.4	562	22	AAW41472	Human polypeptide
45	248	21.3	325	21	AAU94345	Human cell surface

ALIGNMENTS

RESULT 1  
AAU00479  
ID AAU00479 standard; Protein; 209 AA.

AC AAU00479;

DT 04-JUL-2001 (first entry)

XX Human TANGO 405 protein.

DE Human; TANGO 210; clone jthLa152h06; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder.

XX Homo sapiens.

EH Key Location/Qualifiers  
FT Peptide 1..48  
FT /label= Signal\_peptide  
FT /note= "Specifically claimed"

FT Protein 49..209  
FT /label= Mature\_TANGO\_405\_protein  
FT /note= "Specifically claimed"

XX WO200118016-A1.

XX 15-MAR-2001.

XX 30-JUN-2000; 2000WO-US18174.

XX 10-SEP-1999; 99US-0393996.

```
XX (MILL-) MILLENNIUM PHARM INC.
PA Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;
PI WPI; 2001-183280/18.
XX N-PSDB; AAS01375.
DR Isolated nucleic acid molecules encoding proteins useful as modulating
PT agents in regulating a variety of cellular processes are used for
PT treating e.g. cancer and autoimmune disorders -
XX Claim 9; Fig 6A-6C; 326pp; English.
XX The present sequence representing human TANGO 405 protein is isolated
CC from cDNA clone jthual52h06 from a human mixed lymphocyte reaction cDNA
CC library. It is 1 of 6 novel human proteins which include TANGO 210
CC (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394
CC (AAU00473) and INTERCEPT 400 (AAU00476). Novel sequences for murine
CC TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and
CC a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic
CC acids encoding these novel proteins are useful as modulating agents in
CC regulating a variety of cellular processes and can be used to express
CC the proteins in a host cell in gene therapy applications. Human and
CC murine TANGO 405 proteins show sequence homology to murine Dectin-2.
CC TANGO 405 modulates growth, proliferation, survival, differentiation,
CC activity, morphology and movement/migration of human lymphocytes and
CC bone marrow cells and tissues and can be used to prevent, diagnose or
CC treat leukaemia, lymphomas and autoimmune disorders.
XX SQ Sequence 209 AA:
Query Match 100.0%; Score 1165; DB 22; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.4e-111;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMDEQQPQSTKRGWLSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELHSY 60
DB 1 MMDEQQPQSTKRGWLSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELHSY 60
QY 61 HSSLTFCSEGTQKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVFNTE 120
DB 61 HSSLTFCSEGTQKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVFNTE 120
QY 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYEKNVRFWHLGEPNHSAEQACASIV 180
DB 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYEKNVRFWHLGEPNHSAEQACASIV 180
QY 181 FWKPTGWNNDVICETRRNSICEMNKIYL 209
DB 181 FWKPTGWNNDVICETRRNSICEMNKIYL 209
RESULT 2
ID AAY27446
AC AAY27446 standard; Protein; 162 AA.
XX AAY27446;
XX 26-NOV-1999 (first entry)
DT Human SDCMP3 polypeptide.
DE
XX Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;
XX Schering dendritic cell membrane protein; dendritic cell physiology;
XX genetic fingerprinting; cancer immunotherapy; abnormal proliferation;
XX cancer; forensic; human; lectin 73.
XX Homo sapiens.
XX WC09947673-A2.
XX 23-SRP-1999.
PD
```

```
XX 16-MAR-1999; 99WO-US03740.
XX 17-MAR-1998; 98US-0040111.
XX (SCHE ) SCHERING CORP.
XX Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;
XX Philipps JH;
XX WPI; 1999-562114/47.
XX N-PSDB; AAZ07531.
XX Binding compound specific for primate or rodent Schering dendritic cell
PT membrane proteins -
PT Claim 2; Page 80; 89pp; English.
XX The invention relates to a binding compound comprising an antibody
CC binding site which specifically binds to: (a) a primate or rodent
CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a
CC primate SDCMP4 protein. The binding compound can be used to modulate
CC dendritic cell physiology or function. The sequences may also be used as
CC probes in forensic techniques, such as genetic fingerprinting. They can
CC also be used to distinguish tissue and cell types in situ or in vitro.
CC The sequences may also be useful in cancer immunotherapy and in the
CC treatment of conditions associated with abnormal physiology or
CC development, including abnormal proliferation, e.g. cancerous conditions,
CC or degenerative conditions. The present sequence represents the human
CC SDCMP3 polypeptide (lectin 73).
XX SQ Sequence 162 AA:
Query Match 76.0%; Score 885; DB 20; Length 162;
Best Local Similarity 100.0%; Pred. No. 9.4e-83;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMDEQQPQSTKRGWLSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELHSY 60
DB 1 MMDEQQPQSTKRGWLSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELHSY 60
QY 61 HSSLTFCSEGTQKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVFNTE 120
DB 61 HSSLTFCSEGTQKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVFNTE 120
QY 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYEKNVR 162
DB 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYEKNVR 162
RESULT 3
ID AAW63010
XX AAW63010 standard; Protein; 209 AA.
XX AAW63010;
XX 09-NOV-1998 (first entry)
DT Mouse Dectin-2.
DE
XX Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;
XX allergy; autoimmune disease; therapy; vaccine; diagnosis;
XX drug screening.
XX Mus sp.
XX Key Location/Qualifiers
FH 1..13 /label= Cytoplasmic
FT Domain 14..42 /label= Transmembrane
FT Domain 43..209 /label= Extracellular
FT
```



## RESULT 5

04-JUL-2001 (first entry)

Murine; TANGO 210; clone jtm025a11; TANGO 364; TANGO 366; dectin-2; INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator; gene therapy; growth modulator; proliferation; cell differentiation; lymphocyte; bone marrow cell migration; leukaemia; lymphoma; autoimmune disorder.

XX  
PN WO200118016-A1.XX  
PF 30-JUN-2000:

FR 10-SEP-1999, 9903-0393990.  
XX  
XX  
XX

PPI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ  
XX

Isolated nucleic

agents in regulating a variety of cellular processes are used for treating e.g. cancer and autoimmune disorders -  
XX  
XX  
PS Disclosure; Fig 6N-6P; 326pp; English.

The present sequence representing murine TANGO 405 alternative splice variant protein is isolated from cDNA clone jtmMa025all from a long-term bone marrow cDNA library. TANGO 405 (AAU000480) is 1 of 3 novel murine proteins which include TANGO 210 (AAU000470) and INTERCEPT 400 (AAU000477). Six novel human proteins which include TANGO 210 (AAU000469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids encoding these novel proteins are useful as modulating agents in regulating a variety of cellular processes and can be used to express the proteins in a host cell in gene therapy applications. Human and murine TANGO 405 proteins show sequence homology to murine dectin-2. TANGO 405 modulates growth, proliferation, survival, differentiation, activity, morphology and movement/migration of human lymphocytes and bone marrow cells and tissues and can be used to prevent, diagnose or treat leukaemia, lymphomas and autoimmune disorders.

XX	Sequence	209 AA;
SQ		

Query Match	68.1%;	Score 793;	DB 22;	Length 209;
Best Local Similarity	68.1%;	Pred. No. 3.6e-73;		
Matches 145.	Conservative	19.	Mismatches 41.	Indels 8

1960-1961

1962-1963

1964-1965

1966-1967

1968-1969

1970-1971

1972-1973

1974-1975

1976-1977

1978-1979

1980-1981

1982-1983

1984-1985

1986-1987

1988-1989

1990-1991

1992-1993

1994-1995

1996-1997

1998-1999

2000-2001

2002-2003

2004-2005

2006-2007

2008-2009

2010-2011

2012-2013

2014-2015

2016-2017

2018-2019

2020-2021

2022-2023

2024-2025

2026-2027

2028-2029

2030-2031

2032-2033

2034-2035

2036-2037

2038-2039

2040-2041

2042-2043

2044-2045

2046-2047

2048-2049

2050-2051

2052-2053

2054-2055

2056-2057

2058-2059

2060-2061

2062-2063

2064-2065

2066-2067

2068-2069

2070-2071

2072-2073

2074-2075

2076-2077

2078-2079

2080-2081

2082-2083

2084-2085

2086-2087

2088-2089

2090-2091

2092-2093

2094-2095

2096-2097

2098-2099

2100-2101

2102-2103

2104-2105

2106-2107

2108-2109

2110-2111

2112-2113

2114-2115

2116-2117

2118-2119

2120-2121

2122-2123

2124-2125

2126-2127

2128-2129

2130-2131

2132-2133

2134-2135

2136-2137

2138-2139

2140-2141

2142-2143

2144-2145

2146-2147

2148-2149

2150-2151

2152-2153

2154-2155

2156-2157

2158-2159

2160-2161

2162-2163

2164-2165

2166-2167

2168-2169

2170-2171

2172-2173

2174-2175

2176-2177

2178-2179

2180-2181

2182-2183

2184-2185

2186-2187

2188-2189

2190-2191

2192-2193

2194-2195

2196-2197

2198-2199

2200-2201

2202-2203

2204-2205

2206-2207

2208-2209

2210-2211

2212-2213

2214-2215

2216-2217

2218-2219

2220-2221

2222-2223

2224-2225

2226-2227

2228-2229

2230-2231

2232-2233

2234-2235

2236-2237

2238-2239

2240-2241

2242-2243

2244-2245

2246-2247

2248-2249

2250-2251

2252-2253

2254-2255

2256-2257

2258-2259

2260-2261

2262-2263

2264-2265

2266-2267

2268-2269

2270-2271

2272-2273

2274-2275

2276-2277

2278-2279

2280-2281

2282-2283

2284-2285

2286-2287

2288-2289

2290-2291

2292-2293

2294-2295

2296-2297

2298-2299

2300-2301

2302-2303

2304-2305

2306-2307

2308-2309

2310-2311

2312-2313

2314-2315

2316-2317

2318-2319

2320-2321

2322-2323

2324-2325

2326-2327

2328-2329

2330-2331

2332-2333

2334-2335

2336-2337

2338-2339

2340-2341

2342-2343

2344-2345

2346-2347

2348-2349

2350-2351

2352-2353

2354-2355

2356-2357

2358-2359

2360-2361

2362-2363

2364-2365

2366-2367

2368-2369

2370-2371

2372-2373

2374-2375

2376-2377

2378-2379

2380-2381

2382-2383

2384-2385

2386-2387

2388-2389

2390-2391

2392-2393

2394-2395

2396-2397

2398-2399

2400-2401

2402-2403

2404-2405

2406-2407

2408-2409

2410-2411

2412-2413

2414-2415

2416-2417

2418-2419

2420-2421

2422-2423

2424-2425

2426-2427

2428-2429

2430-2431

2432-2433

2434-2435

2436-2437

2438-2439

2440-2441

2442-2443

2444-2445

2446-2447

2448-2449

2450-2451

2452-2453

2454-2455

2456-2457

2458-2459

2460-2461

2462-2463

2464-2465

2466-2467

2468-2469

2470-2471

2472-2473

2474-2475

2476-2477

2478-2479

2480-2481

2482-2483

2484-2485

2486-2487

2488-2489

2490-2491

2492-2493

2494-2495

2496-2497

2498-2499

2500-2501

2502-2503

2504-2505

2506-2507

2508-2509

2510-2511

2512-2513

2514-2515

2516-2517

2518-2519

2520-2521

2522-2523

2524-2525

2526-2527

2528-2529

2530-2531

2532-2533

2534-2535

2536-2537

2538-2539

2540-2541

2542-2543

2544-2

Db 1 MVQERQSQG ~ - - KGVCW - ILRLWSAAVISMLLLSTCFIASCVVTYQFIMDQPSRRLYELH 56

Qy 59 SYHSLTCFSEGTKV--PAWGCPASWKSGSSCYFISSEEKVMKSEQNCVEMGAHLV 116

Db 57 TYHSLTCFSEGTVMSEKMWGCCPNHWKSFSGSSCYLISTKENFWSTSEQNCVQMGAILV 116



```

QY 103 SEQNCVEMGAHLVVFENTEAEQNFIVQOLNESFSYFLGLSDPGQNNWQIDKTPYEKNVR 162
    |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 SEQNCVQMGHVLVVFINTAEQNFITQOLNESLSYFLGLSDPGQNGKQWIDDTFFSQNVR 120
    |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 FWHLGEFNHSAEQCASTVFWKPTGWNNDVICETRRNSICEMKKIYL 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 FWHPEHNPBPERCVSIVYWNPSKMGWNDVFCDSKHNSICEMKKIYL 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AAW63018
ID AAW63018 standard; Protein; 175 AA.
XX
AC AAW63018;
XX
DT 09-NOV-1998 (first entry)
XX
DE Mouse dectin-2 isoform T-gamma.
XX
KW Dectin-2; dendritic cell; C-type lectin; mouse; immunity; adjuvant;
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;
KW drug screening.
XX
OS Mus sp.
XX
PN WO9828332-A2.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-US23761.
XX
PR 20-DEC-1996; 96US-0772440.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Ariizumi K, Takashima A;
XX
WPI; 1998-377594/32.
XX
PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1
PT and -2 - useful, e.g. to regulate immune response, as vaccine
PT adjuvants, for diagnosis and drug screening
XX
PS Claim 10; Page 149; 200pp; English.
XX
CC This polypeptide comprises a truncated isoform, designated T-gamma,
CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)
CC specific member of the C-type lectin family that is essential for
CC DC-mediated T cell activation. It contains a 34-amino acid
CC deletion primarily in the connecting domain compared with the
CC full-length (alpha) dectin-2 polypeptide. It is contemplated that
CC DC in normal tissues express mRNA for different naturally-occurring
CC isoforms of dectin-2. The invention provides: dectin-1 and
CC dectin-2 polypeptides (see also AAW63009-22 and AAW69236-37), useful
CC for purifying T cells, for detecting autoantibodies and for
CC up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see
CC AAW42548-54, AAW42558-60 and AAW44850-51); expression vectors;
CC recombinant host cells; probes and primers; antibodies; compounds
CC that modulate dectin-mediated activation of T cells; transgenic
CC animals useful for studying dectin function and for drug screening;
CC and dectin ligands. Dectin expression can be downregulated for
CC treating allergy and autoimmune disease.
XX
SQ Sequence 175 AA;
    Query Match 56.6%; Score 659; DB 19; Length 175;
    Best Local Similarity 58.3%; Pred. No. 1 6e-59;
    Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;

QY 1 MMEQEQPQSTTEKRG--WLSLRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELH 58
    |::|:| | | | |::| | | |::| | | |::| | | |::| | | |::| | | |::| | | |::|
Db 1 MVQERQSQG---KGVCW-TLRLWSAAVISMLLLSTCFIASCV----- 38

```

```

QY 59 SYHSSILTCFSEGTKVPAMGCCPASMKSGSSCYFISSEKVMKSEQNCVEMGAHLVVFEN 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 -----BKMGGCCPNHHKSGSSCYLLISTKFNWSTSEQNCVQMGHVLVVFEN 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 TEAEQNFIVQOLNESFSYFLGLSDPGQNNWQIDKTPYEKNVRFWHLGEFNHSAEQCAS 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 TEAEQNFITQOLNESLSYFLGLSDPGQNGKQWIDDTFFSQNVRFWHPHPEHNPBPERCVS 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 IVFWKPTGWNNDVICETRRNSICEMKKIYL 209
    |||:| |::| ||||| ||::| ||||| |||||
Db 145 IVYWNPSKMGWNDVFCDSKHNSICEMKKIYL 175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
AAB84215
ID AAB84215 standard; Protein; 213 AA.
XX
AC AAB84215;
XX
DT 06-AUG-2001 (first entry)
XX
DE Amino acid sequence of BDCA-2 antigen isoform.
XX
KW BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;
KW viral infection; autoimmune disease; allergic response; cancer.
XX
OS Homo sapiens.
XX
PN WO200136487-A2.
XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-IB01832.
XX
PR 15-NOV-1999; 99US-0165555.
PR 23-NOV-1999; 99US-0167076.
PR 28-JAN-2000; 2000US-0179003.
PR 07-FEB-2000; 2000US-0180775.
PR 11-APR-2000; 2000US-0196824.
PR 13-APR-2000; 2000US-0197205.
XX
PA (MILT-) MILTENYI BIOTECH GMBH.
XX
PI Schmitz J, Dziunek A, Buck DW;
XX
WPI; 2001-355622/37.
DR N-PSDB; AAF90241.
XX
CC Compositions and cell populations enriched in dendritic cells through
CC use of antigen-binding fragments specific for BDCA-2, BDCA-3 or
CC BDCA-4, are used to treat viral infections, autoimmune diseases,
CC allergic responses and cancer -
XX
PS Claim 122; Fig 5; 115pp; English.
XX
CC The present sequence represents an antigen designated BDCA-2. The
CC specification describes compositions and hematopoietic cell populations
CC enriched in dendritic cells (DCs). These compositions are produced using
CC antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4. The DCs
CC obtained are used to treat viral infections, autoimmune disease, allergic
CC response, and cancer. BDCA-1, BDCA-2, BDCA-3 and BDCA-4 monoclonal
CC antibodies and their antigen-binding fragments are used to detect,
CC enumerate and isolate DC populations from leukapheresis material, whole
CC blood and tonsils and from non-hematopoietic and hematopoietic tissues.
XX
SQ Sequence 213 AA;
    Query Match 54.7%; Score 637; DB 22; Length 213;
    Best Local Similarity 54.0%; Pred. No. 3 8e-57;
    Matches 115; Conservative 30; Mismatches 64; Indels 4; Gaps 2;

QY 1 MMEQEQPQSTTEK-RGWLRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSEL-- 57
    |::|:| | | | |::| | | |::| | | |::| | | |::| | | |::| | | |::| | | |::|

```

Db 1 MVBEPQDREKGLWFWQLKWSMAVVSILLSVCTTVSSVPHNFMWKTVKRLSKLRE 60

QY 58 -HSYHSLTFCSEGTKVPAGCCPASWKSFGSCYFISSEKVKSEQNCVEMGAHLVW 116

Db 61 YQYHPSLTCTVMEGKDIEDWSCCPTPTWTSFQSSCYFISTGMSQWTKSOKNCSVMGADLVV 120

QY 117 FNTAEQNFIVQQLNESFYFLGLSDPGQNNWQIDKTPYBKNVRFWHLGEPNHSAEQC 176

Db 121 INTREQDFTIQNLKRNSYFLGLSDPGGRRHWQVDQTPYNNVTFWHSGEPNLDERC 180

QY 177 ASIVFNKPTGWNVDVICTRRNSICENMKIYL 209

Db 181 AINFRSSEEWGNDIHCHVPOKSICKMKKIYI 213

RESULT 9

AAE12079

ID AAE12079 standard; Protein; 218 AA.

XX AAE12079;

AC AAE12079;

XX 18-DEC-2001 (first entry)

XX Dendritic cell (DC) DCLEC protein.

XX Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;

KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;

KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;

KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;

KW transplant rejection; chronic inflammatory disease; anti-HIV.

XX Unidentified.

OS

XX Key Location/Qualifiers

FX 26..46

FT Domain /note= "Single transmembrane region"

FT Domain 114..211

FT /note= "Extracellular C-type lectin domain"

XX

XX WO200172773-A2.

XX

XX 04-OCT-2001.

XX

XX 28-MAR-2001; 2001WO-EP03542.

XX

XX 29-MAR-2000; 2000US-192934P.

PR 18-MAY-2000; 2000US-205026P.

PR 18-MAY-2000; 2000US-205026P.

PR 19-MAY-2000; 2000US-205767P.

PR 19-MAY-2000; 2000US-205769P.

XX

XX (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;

XX

XX WPI; 2001-616466/71.

DR N-PSDB; AAD19729.

XX

XX New polypeptides for screening therapeutic agonists and antagonists

PT comprise dendritic cell polypeptides -

PT

XX Claim 2; Page 51; 52pp; English.

XX

XX The invention relates to dendritic cell (DC) proteins and their

CC corresponding DNA molecules. A pharmaceutical composition comprising

CC agonist and antagonist of DC proteins are useful for treating abnormal

CC conditions related to both an excess of and insufficient level of

CC expression of DC gene, or related to both an excess of and insufficient

CC activity of DC protein. Soluble form of DC proteins are used as an active

CC ingredient in combination with pharmaceutical acceptable carriers.

CC DC genes and proteins are useful for treating chronic inflammatory

CC diseases, autoimmune diseases, transplant rejection crisis, including

CC inflammatory skin diseases such as contact hypersensitivity, atopic

CC dermatitis or virally-induced immune suppression such as AIDS and cancer.

CC DC protein is useful for inducing immunological response in a mammal, and

CC as immunogen to produce antibodies immunospecific for the polypeptide.

CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic

CC reagent, and for chromosomal identification. The present sequence is

CC dendritic cell (DC) DCLEC protein which is found to belong to the family

CC of C-type lectins with one single carbohydrate recognition domain at the

CC C-terminal end.

XX

SQ Sequence 218 AA;

Query Match 54.1%; Score 630; DB 22; Length 218;

Best Local Similarity 54.5%; Pred. No. 2.1e-56;

Matches 114; Conservative 28; Mismatches 63; Indels 4; Gaps 2;

QY 5 QOQOSTEK-RGWSLRWSVAGISIALLSACFIVSCVVTYHTYGETGKRLSEL---HSY 60

Db 10 QEPQDREKGLWFWQLKWSMAVVSILLSVCTTVSSVPHNFMWKTVKRLSKLREYQQY 69

QY 61 HSSLTFCSEGTKVPAGCCPASWKSFGSCYFISSEKVKSEQNCVEMGAHLVVENTE 120

Db 70 HPSLTCTVMEGKDIEDWSCCPTPTWTSFQSSCYFISTGMSQWTKSOKNCSVMGADLVVINTR 129

QY 121 AEQNFIVQQLNESFYFLGLSDPGQNNWQIDKTPYBKNVRFWHLGEPNHSAPQCASTY 180

Db 130 EQQDFIQLKRNSYFLGLSDPGGRRHWQVDQTPYNNVTFWHSGEPNLDERCAIN 189

QY 181 FWKPTGWNVDVICTRRNSICENMKIYL 209

Db 190 FRSEEWGNDIHCHVPOKSICKMKKIYI 218

RESULT 10

AAE12080

ID AAE12080 standard; Protein; 211 AA.

XX AAE12080;

XX 18-DEC-2001 (first entry)

XX Dendritic cell (DC) DCLEC/SPLICE 1 variant protein.

XX Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;

KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;

KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;

KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;

KW transplant rejection; chronic inflammatory disease; anti-HIV; variant.

XX Unidentified.

OS

XX Key Location/Qualifiers

FX Misc-difference 1 /note= "Encoded by C"

FT

XX WO200172773-A2.

XX

XX 04-OCT-2001.

XX

XX 28-MAR-2001; 2001WO-EP03542.

XX

XX 29-MAR-2000; 2000US-192934P.

PR 18-MAY-2000; 2000US-205026P.

PR 18-MAY-2000; 2000US-205026P.

PR 19-MAY-2000; 2000US-205767P.

PR 19-MAY-2000; 2000US-205769P.

XX

XX (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;

XX

XX WPI; 2001-616466/71.

DR

DR N-PSDB; AAD19730.  
XX New polypeptides for screening therapeutic agonists and antagonists  
PT comprise dendritic cell polypeptides -  
XX  
PS Claim 2; Page 52; 52pp; English.  
XX  
CC The invention relates to dendritic cell (DC) proteins and their  
CC corresponding DNA molecules. A pharmaceutical composition comprising  
CC agonist and antagonist of DC proteins are useful for treating abnormal  
CC conditions related to both an excess of and insufficient level of  
CC expression of DC gene, or related to both an excess of and insufficient  
CC activity of DC protein. Soluble form of DC proteins are used as an active  
CC ingredient in combination with pharmaceutical acceptable carriers.  
CC DC genes and proteins are useful for treating chronic inflammatory  
CC diseases, autoimmune diseases, transplant rejection crisis, including  
CC inflammatory skin diseases such as contact hypersensitivity, atopic  
CC dermatitis or virally-induced immune suppression such as AIDS and cancer.  
CC DC protein is useful for inducing immunological response in a mammal, and  
CC as immunogen to produce antibodies immunospecific for the polypeptide.  
CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
CC reagent, and for chromosomal identification. The present sequence is  
CC dendritic cell (DC) DCLEC/SPLICE 1 variant protein which is found to  
CC belong to the family of C-type lectins with one single carbohydrate  
CC recognition domain at the C-terminal end.  
XX  
SQ Sequence 211 AA;  
Query Match 53.0%; Score 617; DB 22; Length 211;  
Best Local Similarity 52.9%; Pred. No. 4.3e-55;  
Matches 111; Conservative 30; Mismatches 63; Indels 6; Gaps 2;  
QY 6 OPOSTEKR---GWSLRLWSVAGISIALLSACFTVSCVYVYHFTYGTGRRLSEL---HS 59  
DB 2 RPAEDREKGLWFWQLKWSNAVVISLLSVCFVSSVPHNFMYSKTVRLSKRLREYQQ 61  
QY 60 YHSSLTCFSECTKVPAGWCCPASWKSFGSCCYFTISSEKVMKSEQNCVEMGAHLVVNT 119  
DB 62 YHSLTCVMEGKDIEDWSCCPTPTWTSFQSSCYFISTGMSQKSWKSNCSVMGADLVVINT 121  
QY 120 EAEQNFIVQOLNESFYLGLSDPQGNNNQWIDKTPYEKNVRFWHLGEPNHSAEQCAI 179  
DB 122 REEQDFITONLKRNSYFLGLSDPGRRRHQWVDQTPYENVTFWHSGEPNLDERCAII 181  
QY 180 VFWRPTGWNVDVICETRRNSICEMNKIYL 209  
DB 182 NFRSEEWGWNDIHCHVPQKSIKMKKIYI 211  
RESULT 11  
AAW63020  
ID AAW63020 standard; Protein; 168 AA.  
XX  
AC AAW63020;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Mouse dectin-2 isoform T-epsilon.  
XX  
KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;  
KW drug screening.  
XX  
OS Mus sp.  
XX  
PN WO9828332-A2.  
XX  
PD 02-JUL-1998.  
XX  
PF 22-DEC-1997; 97WO-US23761.  
XX  
PR 20-DEC-1996; 96US-0772440.  
XX

(TEXA ) UNIV TEXAS SYSTEM.  
XX  
PI Ariizumi K, Takashima A;  
XX  
DR WPI; 1998-377594/32.  
XX  
PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
PT and -2 - useful, e.g. to regulate immune response, as vaccine  
PT adjuvants, for diagnosis and drug screening  
XX  
PS Claim 12; Page 151; 200pp; English.  
XX  
CC This polypeptide comprises a truncated isoform, designated  
CC T-epsilon, of dectin-2 (see AAW63010), a novel murine dendritic cell  
CC (DC) specific member of the C-type lectin family that is essential  
CC for DC-mediated T cell activation. T-epsilon contains a 43-amino  
CC amino acid deletion in the carbohydrate recognition domain compared  
CC with the full-length (alpha) dectin-2 polypeptide. It is  
CC contemplated that DC in normal tissues express mRNA for different  
CC naturally-occurring isoforms of dectin-2. The invention provides:  
CC dectin-1 and dectin-2 polypeptides (see also AAW63009-22 and  
CC AAW69236-37), useful for purifying T cells, for detecting  
CC autoantibodies and for up-regulating immunity e.g. as vaccine  
CC adjuvants; dectin DNA (see AAW42548-54, AAW42558-60 and AAW44850-51);  
CC expression vectors; recombinant host cells; probes and primers;  
CC antibodies; compounds that modulate dectin-mediated activation of T  
CC cells; transgenic animals useful for studying dectin function and  
CC for drug screening; and dectin ligands. Dectin expression can be  
CC downregulated for treating allergy and autoimmune disease.  
XX  
SQ Sequence 168 AA;  
Query Match 50.9%; Score 592.5; DB 19; Length 168;  
Best Local Similarity 55.4%; Pred. No. 1e-52; Mismatches 16; Indels 49; Gaps 5;  
Matches 118; Conservative 16;  
QY 1 MMQEQOQPQSTEKRG--WLSRLWSVAGISIALLSACFTVSCVYVYHFTYGTGRRLSELH 58  
DB 1 MVQERSQG--KGVCW-TLRWSAAVISMLLLSTCFIASCVVTYQFINDQPSRRLYELH 56  
QY 59 SYHSSLTCFSECTKTV--PAWGCCPASWKSFGSCCYFTISSEKVMKSEQNCVEMGAHLVV 116  
DB 57 TVHSSLTCFSECTVMSEKMWGCCPNHWKSGSCCYLISTKENFWSTSEQNCVQMGHLVV 116  
QY 117 FNTAEQNFIVQOLNESFYLGLSDPQGNNNQWIDKTPYEKNVRFWHLGEPNHSAEQC 176  
DB 117 INTEAEQNFITQOLNESLSYFLGLSN----- 142  
QY 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209  
DB 143 -----PSKWGWNVDVFCDSKHSICEMKKIYL 168  
RESULT 12  
AAW63017  
ID AAW63017 standard; Protein; 145 AA.  
XX  
AC AAW63017;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Mouse dectin-2 isoform T-beta.  
XX  
KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;  
KW drug screening.  
XX  
OS Mus sp.  
XX  
PN WO9828332-A2.  
XX  
PD 02-JUL-1998.  
XX

PF 22-DEC-1997; 97WO-US23761.  
PR 20-DEC-1996; 96US-0772440.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PA Arizumi K, Takashima A;  
XX WPI; 1998-377594/32.  
XX  
DR Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
PT and -2 - useful, e.g. to regulate immune response, as vaccine  
PT adjuvants, for diagnosis and drug screening  
XX  
PS Claim 9; Page 148-149; 200pp; English.  
XX  
CC This polypeptide comprises a truncated isoform, designated T-beta,  
CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)  
CC specific member of the C-type lectin family that is essential for  
CC DC-mediated T cell activation. It contains a 65-amino acid  
CC deletion primarily in the transmembrane and connecting domains  
CC compared with the full-length (alpha) dectin-2 polypeptide. It is  
CC contemplated that DC in normal tissues express mRNA for different  
CC naturally-occurring isoforms of dectin-2. The invention provides:  
CC dectin-1 and dectin-2 polypeptides (see also AAW63009-22 and  
CC AAW63036-37), useful for purifying T cells, for detecting  
CC autoantibodies and for up-regulating immunity e.g. as vaccine  
CC adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and AAV44850-51);  
CC expression vectors; recombinant host cells; probes and primers;  
CC antibodies; compounds that modulate dectin-mediated activation of  
CC T cells; transgenic animals useful for studying dectin function  
CC and for drug screening; and dectin ligands. Dectin expression can  
CC be downregulated for treating allergy and autoimmune disease.  
XX  
SQ Sequence 145 AA;  
  
Query Match 50.7%; Score 591; DB 19; Length 145;  
Best Local Similarity 73.9%; Pred. No. 1.2e-52;  
Matches 99; Conservative 11; Mismatches 24; Indels 0; Gaps 0;  
  
Qy 76 WGCCPASWKSFGSSCYFISSEERKWSKSRQNCVEMGAHLVFTFAQNFIVQQLNESFS 135  
Db 12 WGCCPNHWKSGSSCYLSTKENFWSSTQNCVOMGAHLVINTAEQNFITQQLNESLS 71  
  
Qy 136 YFLGLSDPOGNNWOWIDTPTYEKNVRFWHLGEPNHSABQACASIVFWKPTGKWNDVICE 195  
Db 72 YFLGLSDPOGNGKWQWIDTPTSQNVRFWHPHPNLPPEERCVSIVVWNPFSKGNWDFCD 131  
  
Qy 196 TRNSICEMKKIYL 209  
Db 132 SKHNSICEMKKIYL 145  
  
RESULT 13  
AAU00480  
ID AAU00480 standard; Protein; 178 AA.  
AC AAU00480;  
XX  
XX 04-JUL-2001 (first entry)  
DE Murine TANGO 405 protein.  
XX  
KW Murine; TANGO 210; clone jtmMa025a11; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder.  
OS Mus sp.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..42

FT  
FT Protein /label= Signal\_peptide  
FT 43..178  
XX /label= Mature\_TANGO\_405\_protein  
XX  
PN WO200118016-A1.  
XX  
XX 15-MAR-2001.  
XX  
XX 30-JUN-2000; 2000WO-US18174.  
XX  
XX 10-SEP-1999; 99US-0393996.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX  
XX WPI; 2001-183280/18.  
XX N-PSDB; AAS01376.  
XX  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
XX agents in regulating a variety of cellular processes are used for  
XX treating e.g. cancer and autoimmune disorders -  
XX  
XX Disclosure; Fig 6E-6F; 326pp; English.  
XX  
XX The present sequence representing murine TANGO 405 protein is isolated  
XX from cDNA clone jtmMa025a11 from a long-term bone marrow CDNA library.  
XX An alternative splice variant of this protein is also described  
XX (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
XX INTERCEPT 400 (AAU00477). Six novel human proteins which include  
XX TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
XX INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
XX (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
XX described. The nucleic acids encoding these novel proteins are useful as  
XX modulating agents in regulating a variety of cellular processes and can  
XX be used to express the proteins in a host cell in gene therapy  
XX applications. Human and murine TANGO 405 proteins show sequence homology  
XX to murine dectin-2. TANGO 405 modulates growth, proliferation, survival,  
XX differentiation, activity, morphology and movement/migration of human  
XX lymphocytes and bone marrow cells and tissues and can be used to prevent,  
XX diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 178 AA;  
  
Query Match 44.7%; Score 521; DB 22; Length 178;  
Best Local Similarity 66.5%; Pred. No. 2.4e-45;  
Matches 105; Conservative 12; Mismatches 31; Indels 10; Gaps 5;  
  
Qy 1 MMQEQQPOSTEKRG--WLSURLNSVAGTISTALLSACFTVSCVVTYHYFTYGETKRLSELH 58  
Db 1 MVQERQSQG---KGVCW-TLRWSAAVISMILLSTCTFIASCVTYQFIMDQPSRRLYELH 56  
  
Qy 59 SYHSSLTCFSGTKV--PAWGCCPASWKSRSQSCYFTSSSEKVKWSKSEQNCVEMGAHLV 116  
Db 57 TYHSSLTCFSGTWTWSEKMMGCCPNHWKSRGSSCYLSTKENFWSSTQNCVOMGAHLV 116  
  
Qy 117 FNTAEQNFIVQQLNESFSYFLGLSDPQ--GNNNWOWI 152  
Db 117 INTEAEQNFITQQLNESLSYFLGLSDPKVWANGNGSMI 154  
  
RESULT 14  
AAW63019  
ID AAW63019 standard; Protein; 134 AA.  
XX  
XX AAW63019;  
XX  
XX 09-NOV-1998 (first entry)  
XX  
XX Mouse dectin-2 isoform T-delta.  
XX  
XX Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;

KW drug screening.  
 XX Mus sp.  
 OS WO9828332-A2.  
 PN  
 XX 02-JUL-1998.  
 PD  
 XX  
 PF 22-DEC-1997; 97WO-US23761.  
 XX  
 PR 20-DEC-1996; 96US-0772440.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Ariizumi K, Takashima A;  
 XX  
 DR WPI; 1998-377594/32.  
 XX  
 XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 PT and -2 - useful, e.g. to regulate immune response, as vaccine  
 PT adjuvants, for diagnosis and drug screening  
 XX  
 PS Claim 10; Page 149-150; 200pp; English.  
 XX  
 CC This polypeptide comprises a truncated isoform, designated T-delta,  
 CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)  
 CC specific member of the C-type lectin family that is essential for  
 CC DC-mediated T cell activation. T-delta contains a 41-amino acid  
 CC deletion in the C-terminus of the carbohydrate recognition domain  
 CC as well as a 34-amino acid deletion primarily in the connecting  
 CC domain compared with the full-length (alpha) dectin-2 polypeptide.  
 CC It is contemplated that DC in normal tissues express mRNA for  
 CC different naturally-occurring isoforms of dectin-2. The invention  
 CC provides: dectin-1 and dectin-2 polypeptides (see also AAW63009-22  
 CC and AAW6236-37), useful for purifying T cells, for detecting  
 CC autoantibodies and for up-regulating immunity e.g. as vaccine  
 CC adjuvants; dectin DNA (see AAW42548-54, AAW42558-60 and AAW44850-51);  
 CC expression vectors; recombinant host cells; probes and primers;  
 CC antibodies; compounds that modulate dectin-mediated activation of T  
 CC cells; transgenic animals useful for studying dectin function and  
 CC for drug screening; and dectin ligands. Dectin expression can be  
 CC downregulated for treating allergy and autoimmune disease.  
 XX  
 SO Sequence 134 AA;  
 Query Match 42.3%; Score 493; DB 19; Length 134;  
 Best Local Similarity 57.5%; Pred. No. 1.3e-42;  
 Matches 96; Conservative 10; Mismatches 23; Indels 38; Gaps 4;  
 Qy 1 MMQEQPQSTKRG--WLSRLNSVAGISIALLSACFIVSCVVTYHFTYGETGKRISLH 58  
 Db 1 MVOERQSG---KGCW-TLRLWSAAVISMILLSTCFIASCV----- 38  
 Qy 59 SYHSLTCFSECTKVPWAGCCPASWKSFGSSCYFISSEKVKSEONCVEMGAHLVFN 118  
 Db 39 -----EKWVGCCPNHWKSGSSCYLISTKENFSTSEQNCVQMG AHLVIN 84  
 Qy 119 TEAQNFITVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWH 165  
 Db 85 TEAQNFITVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWH 131  
 RESULT 15  
 AAU98015  
 ID AAU98015 standard; Protein; 148 AA.  
 XX  
 AC AAU98015;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human partial dendritic cell immunoreceptor #2.  
 XX  
 KW Human; dendritic cell immunoreceptor; cytostatic;

KW antiasthmatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;  
 KW asthma; inflammation; obesity; diabetes; central nervous system disorder;  
 KW Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;  
 KW cardiovascular disorder; myocardial infarction; ischaemic heart disease;  
 KW congestive heart failure; chronic obstructive pulmonary disease; COPD.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200232958-A2.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-BP11812.  
 XX  
 PR 16-OCT-2000; 2000US-240096P.  
 PR 27-AUG-2001; 2001US-314661P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 XX Smolyar A;  
 PI  
 XX  
 DR WPI; 2002-463308/49.  
 DR N-PSDB; ABX52914.  
 XX  
 XX Novel human dendritic cell immunoreceptor polypeptide, useful for  
 PT treating cancer, asthma, obesity, diabetes, central nervous system  
 PT disorder or cardiovascular disorder -  
 XX  
 PS Claim 25; Fig 16; 114pp; English.  
 XX  
 CC The invention relates to a purified human dendritic cell immunoreceptor  
 CC polypeptide and its encoding nucleic acid. Also included are an  
 CC expression vector comprising the nucleic acid, a host cell containing  
 CC the vector, reducing the activity of human dendritic cell immunoreceptor  
 CC involves contacting a cell with a reagent which specifically binds to  
 CC of the polypeptide or nucleic acid and a reagent that modulates the activity  
 CC of the polypeptide or nucleic acid. The reagent is useful for the  
 CC preparation of a medicament for modulating the activity of human  
 CC dendritic cell immunoreceptor in a disease such as cancer, allergy,  
 CC anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous  
 CC system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or  
 CC dementia), or a cardiovascular disorder (e.g. myocardial infarction,  
 CC ischaemic heart disease, congestive heart failure), chronic obstructive  
 CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
 CC disorders are listed in the specification). The polypeptide and nucleic  
 CC acid are useful for identifying test compounds which act as agonists or  
 CC antagonists, for raising specific antibodies, and as a bait protein in a  
 CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
 CC diagnostic assays for detecting diseases and abnormalities or  
 CC susceptibility to disease and abnormalities related to the presence of  
 CC mutations. The present sequence represents a partial sequence of a  
 CC dendritic cell immunoreceptor.  
 XX  
 SO Sequence 148 AA;  
 Query Match 39.4%; Score 459.5; DB 23; Length 148;  
 Best Local Similarity 54.1%; Pred. No. 3.9e-39;  
 Matches 80; Conservative 20; Mismatches 47; Indels 1; Gaps 1;  
 Qy 63 SLTCSFEGTKV-PAMGCCPASWKSFGSSCYFISSEKVKSEONCVEMGAHLVFNTEA 121  
 Db 1 SVNISSSLIWEWDSCTPTWTSFQSSCYFISTGMQSWTKSKNCSVMGADLVVINTRE 60  
 Qy 122 EQNFIVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWHGEPNHSSEQASIVF 181  
 Db 61 EQDFTIQLNKRSSYFLGLSDPQGNWIDKTPYEKNVREWHGEPNHSSEQASIVF 120  
 Qy 182 WKPTGWGNDVICETERNISCEMNNKIYL 209  
 Db 121 RSSEWGWNDIHCVPQKICKMKKIYI 148

Search completed: February 11, 2003, 22:56:57

Job time : 37 secs



```
QY 59 SYHSLTCFSEGTKV--PAWGCPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVV 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 57 TYSLSLTCFSEGMTVSEKMWGCCPNHWKSGSSCYLISTKENFWSTSEQNCVOMGAHLVV 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 117 FNTAEQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHSABQC 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 INTEAQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFSONRVFHPHPNLPERC 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 177 ASIVFWKPTGWGNDVICETRNNSICEMKKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 177 VSIYWNPSKWGNDVFCDSKHNSICEMKKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 2
US-08-772-440-21
; Sequence 21, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-772-440-21

Query Match 58.3%; Score 679; DB 3; Length 167;
Best Local Similarity 70.7%; Pred. No. 6.7e-65;
Matches 118; Conservative 14; Mismatches 33; Indels 2; Gaps 1;

QY 45 FTYGTEGKRLSELHSHSLTCFSEGTKV--PAWGCPASWKSFGSSCYFISSEKVKWSK 102
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 FIMDQPSRLRYELHTYHSLTCFSEGMTVSEKMWGCCPNHWKSGSSCYLISTKENFWST 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 103 SEQNCVEMGAHLVVFNTEAQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVR 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 SEQNCVOMGAHLVVFNTEAQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFSONRV 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 163 FWHLGPENHSABQCAISIVFWKPTGWGNDVICETRNNSICEMKKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 FWHHPNLPERCVSIVYWNPSKWGNDVFCDSKHNSICEMKKIYL 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 3
```

```
US-08-772-440-15
; Sequence 15, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-772-440-15

Query Match 56.8%; Score 659; DB 3; Length 175;
Best Local Similarity 58.3%; Pred. No. 9.7e-63;
Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;

QY 1 MMQEQQPQSTKRG--WLSRLKWSAGISTALLSACFIVSCVVTYHFTYGETCKRLSELH 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MVQERQSG---KGVCM-TLRLWSAAVISMLLSTCFIASCV----- 38
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 59 SYHSLTCFSEGTKVPAWGCPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVVFN 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 39 -----EKMWGCCPNHWKSGSSCYLISTKENFWSTSEQNCVOMGAHLVVIN 84
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 119 TEAQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHSABQCA 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 85 TEAQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFSONRVFHPHPNLPERCVS 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 179 IVFWKPTGWGNDVICETRNNSICEMKKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 145 IVYWNPSKWGNDVFCDSKHNSICEMKKIYL 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 4
US-08-772-440-17
; Sequence 17, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
```





## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-772-440-23

Query Match 48.5%; Score 565; DB 3; Length 131;

Best Local Similarity 73.3%; Pred. No. 7.2e-53;

Matches 96; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 79 CPASWKSFGSGSCYFISSEKVKWSQNCVEMGAHLVFNTEAEQNFIQQQLNESFSYFL 138

Db 1 CPNHWSFGSGCYLISTKENFNWSTSEQNCVQMGALHVINTEAEQNFIQQQLNESLSYFL 60

QY 139 GLSDPGQNNWQWIDKTPYEKNVRFVHLGEPNHSAEQCAIVFWKPTGQWNDVICETR 198

Db 61 GLSDPGQNGKWQWIDTTPFSQNVRFVHPHEPNLPERCVSIYVWNP SKGWNDVFCDSKH 120

QY 199 NSICEMKKIYL 209

Db 121 NSICEMKKIYL 131

## RESULT 7

US-08-772-440-27

; Sequence 27, Application US/08772440

; Patent No. 6046158

; GENERAL INFORMATION:

; APPLICANT: Ariizumi, Kiyoshi

; APPLICANT: Takashima, Akira

; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE

; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White &amp; Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/772,440

; FILING DATE: CONCURRENTLY HERewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, David L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: UTXD:493

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/418-3000

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-772-440-27

Query Match

Best Local Similarity 48.5%; Score 565; DB 3; Length 131;

Matches 96; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 79 CPASWKSFGSGSCYFISSEKVKWSQNCVEMGAHLVFNTEAEQNFIQQQLNESFSYFL 138

Db 1 CPNHWSFGSGCYLISTKENFNWSTSEQNCVQMGALHVINTEAEQNFIQQQLNESLSYFL 60

QY 139 GLSDPGQNNWQWIDKTPYEKNVRFVHLGEPNHSAEQCAIVFWKPTGQWNDVICETR 198

Db 61 GLSDPGQNGKWQWIDTTPFSQNVRFVHPHEPNLPERCVSIYVWNP SKGWNDVFCDSKH 120

QY 199 NSICEMKKIYL 209

Db 121 NSICEMKKIYL 131

## RESULT 8

US-08-772-440-16

; Sequence 16, Application US/08772440

; Patent No. 6046158

; GENERAL INFORMATION:

; APPLICANT: Ariizumi, Kiyoshi

; APPLICANT: Takashima, Akira

; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE

; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White &amp; Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/772,440

; FILING DATE: CONCURRENTLY HERewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, David L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: UTXD:493

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/418-3000

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 134 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-772-440-16

Query Match 42.3%; Score 493; DB 3; Length 134;

Best Local Similarity 57.5%; Pred. No. 3.6e-45;

Matches 96; Conservative 10; Mismatches 23; Indels 38; Gaps 4;

QY 1 MMQEQOQOSTEKRQ--WLSRLWSVAGISTALLSACFIVSCVVTYHFTYGETGKRLSELH 58

Db 1 MVQERQSGQ---KGVCH-TLRLWSAAVISMILLSTCFIASCV----- 38

QY 59 SYHSLTCTFSEGTQVPWAGCCPASWKSFGSGCYFISSEKVKWSQNCVEMGAHLVFN 118

Db 39 -----EKMWGCCPNHWSFGSGCYLISTKENFNWSTSEQNCVQMGALHVVIN 84

QY 119 TEAFQNFIVQQLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFVH 165

Db 85 TEAQNFITQQLNESLSYFLGLSDPGQNGKWQWIDTTPFSQNVRFVH 131

## RESULT 9

US-09-111-470-2

; Sequence 2, Application US/09111470

; Patent No. 6277959

```

; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-111-470-2

```

```

Query Match 39.3%; Score 458; DB 4; Length 237;
Best Local Similarity 44.4%; Pred. No. 4,3e-41;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

QY 26 ISTALLSACFIVSCVYVYHTFYGTGKRLSELHSHYSSLTGFCSEGTKV--PAWGCCPASW 83
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 IFFLLAISFFIAFVFFQ-KYSQLEKKTKRELVTHTTECVKKNMPVEETAWSCCPKNW 110
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 84 KFGSSCYFISSEKVKWSKSEONCVEMGAHLVVFNTFAFONFIVVOLNESHYFLGLSDP 143
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 KFSSNCYFISTESASQDSKDCARMEAHLLVINTQEEQDFIFONQESAYFVGLSDP 170
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 144 QGNNNQWIDKTPYEKNVRFHGLFEPNHSAEQASIVFWK-PTGNGWMDVICETRRNSIC 202
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 EGRHWQWVDQTPYNESSTFWHPREPSDNERCVVINFKSPKRWGWDVNCGLGQRSVC 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 203 ENNKIYL 209
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 EMKKIHL 237
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
US-09-111-470-8
; Sequence 8, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.

```

```

; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-111-470-8

Query Match 37.4%; Score 436; DB 4; Length 238;
Best Local Similarity 39.9%; Pred. No. 9,8e-39;
Matches 85; Conservative 45; Mismatches 73; Indels 10; Gaps 8;

QY 3 QEQQPQSTKKGWLSLRMSVAGISIALLSACFIVSCVYVYHTFYGE-TGKRLSELHSHYH 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 REKPIRDLRKPGSPSLTSLM-LLLLLLATTFLVAFII-YFQKYSQLEEKKAANKIMH 87
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 SSLTGFSEGTKV--PAWGCCPASWKSFGSSCYFIS--SEEKVKWSKSEONCVEMGAHLVVF 117
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 NELNCTKSVSPMEDKVKWSCCPKDWRLFGSHCYLVTPTVSSASWKNKSEENCSRMGAHLVVI 147
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 118 NTEAEQNFIIVOLNESHYFLGLSDPQGNNNQWIDKTPYEKNVRFHGLFEPNHSABQCA 177
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 QSQEEQDFITGLDTHAAAYFGLWD-TGHRQWQWVDQTPVEESITFWHNGEPSSGNEKCA 206
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 178 SIIVF-WKPTGNGWMDVICETRRNSICENNKIYL 209
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 TLIYRWK-TGNGWMDISCSLQKQSVCOMKKINL 238
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-08-772-440-32
; Sequence 32, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/772,440  
FILING DATE: CONCURRENTLY HERewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UTXD:493  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 85 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-772-440-32

Query Match 28.7%; Score 334.5; DB 3; Length 85;  
Best Local Similarity 67.4%; Pred. No. 1.6e-28;  
Matches 60; Conservative 6; Mismatches 16; Indels 7; Gaps 1;  
QY 58 HSYHSLTCTFSEGTGKVPAGCCPASWKSFGSSCYFTISSEKVMKSEQNCVEMGAHLVVF 117  
DB 2 HHHHGSAC-----ELWGCCPNHWKSGSSCYLIISTKENFWSSTSEQNCVQMGALVVI 54

QY 118 NTEAEQNFIVQQLNESFSYFLGLSDPQGN 146  
DB 55 NTEADENFTQQLNESLSYFLGLSDPQGN 83

RESULT 12  
US-09-111-470-6  
Sequence 6, Application US/09111470  
Patent No. 6277959  
GENERAL INFORMATION:  
APPLICANT: Valladeau, Jenny  
APPLICANT: Ravel, Odile  
APPLICANT: Bates, Elizabeth E.M.  
APPLICANT: Ford, John  
APPLICANT: Saeland, Sem  
APPLICANT: Lebecque, Serge J.E.  
TITLE OF INVENTION: Mammalian Membrane Protein Genes;  
TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,470  
FILING DATE: 08-JUL-1998  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/053,080  
FILING DATE: 09-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: SF0695  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)852-9196  
TELEFAX: (650)496-1200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-111-470-6  
Query Match 21.6%; Score 252; DB 4; Length 287;  
Best Local Similarity 25.6%; Pred. No. 5.7e-19;  
Matches 68; Conservative 40; Mismatches 72; Indels 86; Gaps 9;  
QY 16 LSURLWSVAGISTALLSACFIVSCVYTHYFTYGETCKRLSELHSYHSLTCTFSE----- 69  
DB 30 LAQRLCSM--VCFSLLSALSNILLVVICVTGQSQAQLQAEKLSLKEAFNSFSSLTLEV 87  
QY 70 -----CTKVPAMG----- 77  
DB 88 QAISTHGGSVGDKITSLGAKLEKQOQDLKADHDALLFLHKHFPVDLRFVACQMELLHSNG 147  
QY 78 ----CCPASWKSFGSSCYFTISSEKVMKSEQNCVEMGAHLVVFNTAEQNFIVQQLNES 133  
DB 148 SQTCCPVNVYHVGQSCYFTSHSGKAWAEAKYCOLENAHLLVIVINSWEQKFIQHTN-P 206  
QY 134 PSYFLGLSDPQGNNNWQIDTKPYEKNVRFHILGEPNH-----SAEQCASIVENKPTG 186  
DB 207 FNTWIGLTDSDG--SWKWDGTDYRHNKNAVTPQDNHGHGELGSGDCVEV---QPDG 261  
QY 187 WGND-----VICETRRNSICEM 204  
DB 262 -RWNDPCLQVYRWVCEKRRNATGEV 286

RESULT 13  
US-09-111-470-4  
Sequence 4, Application US/09111470  
Patent No. 6277959  
GENERAL INFORMATION:  
APPLICANT: Valladeau, Jenny  
APPLICANT: Ravel, Odile  
APPLICANT: Bates, Elizabeth E.M.  
APPLICANT: Ford, John  
APPLICANT: Saeland, Sem  
APPLICANT: Lebecque, Serge J.E.  
TITLE OF INVENTION: Mammalian Membrane Protein Genes;  
TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,470  
FILING DATE: 08-JUL-1998

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-111-470-10

Query Match 21.5%; Score 250; DB 4; Length 273;
Best Local Similarity 32.4%; Pred. No. 8.7e-19;
Matches 55; Conservative 28; Mismatches 55; Indels 32; Gaps
;

QY 53 RLSLHSHYSSLTG-----FSECTKYPANGCCPASWKSFGSCYFISSEKVM 100
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 RVQOLVODLKKLCQVATLNNGEEAEETG-----CCPVNWEHQDSCYWFSGMSW 159
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 SKSEQNCVMGAHLVFNTEAEQNFIVQQLNESFSYFLGLSDPQGNNNQWIDKTPYKN 160
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 AEAKYCOLKNAHLVINSREEQNFQYKILGSAYTW-MGLSDEG--AKKWDGTDYATG 216
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 161 VRWHLIGEPNH-----SBEQASIVFWKPTGWNVDVICETRRNSICE 203
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 FQWKKPQPDWQOGLGGEDCA---HFHPDG-RWNDDVCORPYHWVCE 262
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-08-688-342-4
; Sequence 4, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1235724
US-08-688-342-4

```

Query Match 21.5%; Score 250; DB 2; Length 292;  
Best Local Similarity 32.4%; Pred. No. 9.5e-19;  
Matches 55; Conservative 28; Mismatches 55; Indels 32; Gaps 7;

QY	53	RLSELHSHSLTC-----FSEGTKVPAMGCCPASMKSEGGSSCYFISSEK YW	100
Db	125	RVQQLVQDLKKLTCQVATLNNNGEEASTEGT-----CCPVNWVEHQDSCYWFESHGMSW	178
QY	101	SKSEQNCVEMGAHLVVVENTAEQNFIVQQLNFSFSYFLGLSDPQGNNNWQWIDKTPYEKN	160
Db	179	AEAEKYCQLKNAHLVVINSREEQNFQKYLGSAYTW-MGLSDPEG--AKKWVDGTDYATG	235
QY	161	VRFWHLGEPNH-----SAEQCASIVFWKPTGHWGNDVICETRNSICE	203
Db	236	FQWKKFQCPDDWQHGGLGGEDCA---HFHPDG-RWNDVVCQRPYHWVCE	281

Search completed: February 11, 2003, 22:59:48  
Job time : 15 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:59:51 : Search time 11 Seconds  
(without alignments)  
485.429 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTQKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*

6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*

7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*

9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*

10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*

11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*

12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*

13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*

14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	637	54.7	213	12	US-10-090-466-2
2	535.5	46.0	182	12	US-10-090-466-4
3	458	39.3	237	9	US-09-862-802-2
4	458	39.3	237	9	US-09-870-759-49
5	458	39.3	243	10	US-09-764-870-464
6	458	39.3	246	10	US-09-764-870-303
7	436	37.4	238	9	US-09-862-802-8
8	398.5	34.2	219	9	US-09-905-291A-377
9	398.5	34.2	219	9	US-09-965-529-3
10	398.5	34.2	219	9	US-09-902-853-377
11	398.5	34.2	219	9	US-09-907-824-377
12	398.5	34.2	219	9	US-09-907-841-377
13	398.5	34.2	219	9	US-09-904-011-377
14	398.5	34.2	219	9	US-10-174-590-24
15	398.5	34.2	219	9	US-10-176-758-24
16	398.5	34.2	219	9	US-10-175-737-24
17	398.5	34.2	219	9	US-09-906-742-377
18	398.5	34.2	219	9	US-10-173-706-24
19	398.5	34.2	219	9	US-10-175-738-24

20	398.5	34.2	219	9	US-10-175-752-24
21	398.5	34.2	219	9	US-10-176-482-24
22	398.5	34.2	219	9	US-10-176-757-24
23	398.5	34.2	219	9	US-10-176-913-24
24	398.5	34.2	219	9	US-10-180-552-24
25	398.5	34.2	219	9	US-10-180-557-24
26	398.5	34.2	219	9	US-09-906-838-377
27	398.5	34.2	219	9	US-09-907-613-377
28	398.5	34.2	219	9	US-09-907-942-377
29	398.5	34.2	219	9	US-10-173-700-24
30	398.5	34.2	219	9	US-10-174-572-24
31	398.5	34.2	219	9	US-10-174-579-24
32	398.5	34.2	219	9	US-10-174-582-24
33	398.5	34.2	219	9	US-10-174-588-24
34	398.5	34.2	219	9	US-10-175-739-24
35	398.5	34.2	219	9	US-10-175-740-24
36	398.5	34.2	219	9	US-10-175-743-24
37	398.5	34.2	219	9	US-10-176-488-24
38	398.5	34.2	219	9	US-10-176-492-24
39	398.5	34.2	219	9	US-10-176-747-24
40	398.5	34.2	219	9	US-10-176-750-24
41	398.5	34.2	219	9	US-10-176-985-24
42	398.5	34.2	219	9	US-10-176-987-24
43	398.5	34.2	219	9	US-10-176-991-24
44	398.5	34.2	219	9	US-10-176-992-24
45	398.5	34.2	219	9	US-10-176-993-24

ALIGNMENTS

RESULT 1

US-10-090-466-2

; Sequence 2, Application US/10090466

; Patent No. US20020137914A1

; GENERAL INFORMATION:

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Mathur, Brian

; APPLICANT: Cullinan, Emily B.

; TITLE OF INVENTION: No. US20020137914A1 Human Dectin Proteins and Polynucleotide

; FILE REFERENCE: LEX-0315-USA

; CURRENT APPLICATION NUMBER: US/10/090.466

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 60/274,961

; PRIOR FILING DATE: 2001-03-12

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 213

; TYPE: PRT

; ORGANISM: homo sapiens

US-10-090-466-2

Query Match	54.7%	Score 637;	DB 12;	Length 213;
Best Local Similarity	54.0%	Pred. No. 1.2e-56;		
Matches	115;	Conservative	30;	Mismatches 64;
				Indels 4;
				Gaps 2;
Qy	1	MMQEQPQSTQK-RGWLSRLWSVAGISALLSACFTVSCVYVYHFTYGTGKRLSEL--	57	
Db	1	MVPEEPQDREKGLWMFQKWSMAVVSILLSCFTVSSVVPVHFNMYSKTVKRLSKLRE	60	
Qy	58	-HSYHSSILTCFSEGTQKVPAMGCCPASWKSFGSSCYFISSEKVMKSEQNCVEMGAHLV	116	
Db	61	YQYHPLSTCVMEKGDIEDSCCPTPTWTSQSSCYFISTCMQSWTKSOKKNCVYGADLV	120	
Qy	117	FNTAEQNFIVQOLNESFSYFLGLSDPQGNMNMWIDKTPYEKNVRFWHLCEPNHSAEQ	176	
Db	121	INTREODFIQNLKRNSYFLGLSDPGRRHMQWVDQTPYNNVNTFWHSGEPNNLDRC	180	
Qy	177	ASIVFWKPTGWNVDVICETRRNSICEMNKIYL	209	
Db	181	AIINFRSEEWGNDIHCHVPQKSICKMKKIYI	213	

```
RESULT 2
US-10-090-466-4
; Sequence 4, Application US/10090466
; Patent No. US20020137914A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Cullinan, Emily B.
; TITLE OF INVENTION: No. US20020137914A1 Human Dectin Proteins and Polynucleotides
; FILE REFERENCE: LEX-0315-USA
; CURRENT APPLICATION NUMBER: US/10/090.466
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/274,961
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 182
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(182)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-090-466-4

Query Match 46.0%; Score 535.5; DB 12; Length 182;
Best Local Similarity 54.7%; Pred. No. 1.5e-46;
Matches 94; Conservative 23; Mismatches 52; Indels 3; Gaps 1;

Qy 41 VTYHTYGTGKRLSEL---HSYHSLTCTFSGTKVPAGWCCPASWSKSGSCYFISSEE 97
Db 11 VPHFNWYKTVKRLSKLREYQYHXSLTCTVMEGKDIEDWSCCPTPTWTFQSSCYFISIGM 70
Qy 98 KVMKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDPQNNKQWIDTKTPY 157
Db 71 QSWTKSQKNCVAGADLVVINTREEQDFIQNLKRNSSYFLGLSDPGGRRHWQWVDQTPY 130
Qy 158 EKXNVFHLGEPNHSAEQCAIVFWKPTGQWGNVDVICTRRNSICEMNKIYL 209
Db 131 NENVTFHSGEPNLDRCALINFRSSEEGWNDIHCHVPQKSIKMKIYI 182

RESULT 3
US-09-862-802-2
; Sequence 2, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-2

Query Match 39.3%; Score 458; DB 9; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.3e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

Qy 26 ISTALLSACFIVSCVVTYHTYGETGKRLSELHSHYSSLTCTFSEGTGV--PAWGCCPASW 83
Db 52 IFFLLAISFFIAFVIFQ-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSCCPKNW 110
Qy 84 KSGSSCYFISSEKVMKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
Db 111 KFSFSSNCYFISTESASWQDSEKDCARMEALLVINTQEEQDFIQNLQESAFYVGLSDP 170
Qy 144 QGNNNWQIDKTPYEKNVRFHGLGEPNHSAEQCAIVFWK-PTGQWGNVDVICTRRNSIC 202
Db 171 EGOHRWQWVDQTPYNESSTFWHPREPSDPNRCVVLNFRKSPKRWGNDVNCILGPORSVC 230
Qy 203 EMNKIYL 209
Db 231 EMMKIHL 237

RESULT 4
US-09-870-759-49
; Sequence 49, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-49

Query Match 39.3%; Score 458; DB 9; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.3e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

Qy 26 ISTALLSACFIVSCVVTYHTYGETGKRLSELHSHYSSLTCTFSEGTGV--PAWGCCPASW 83
Db 52 IFFLLAISFFIAFVIFQ-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSCCPKNW 110
Qy 84 KSGSSCYFISSEKVMKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
Db 111 KFSFSSNCYFISTESASWQDSEKDCARMEALLVINTQEEQDFIQNLQESAFYVGLSDP 170
Qy 144 QGNNNWQIDKTPYEKNVRFHGLGEPNHSAEQCAIVFWK-PTGQWGNVDVICTRRNSIC 202
Db 171 EGOHRWQWVDQTPYNESSTFWHPREPSDPNRCVVLNFRKSPKRWGNDVNCILGPORSVC 230
Qy 203 EMNKIYL 209
Db 231 EMMKIHL 237

RESULT 5
US-09-764-870-464
; Sequence 464, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PYZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-464
```





Wed Feb 19 09:55:49 2003

us-09-766-511b-53.rspt

Page 7

Job time : 30 secs

```

      : | | : : | | | | : | | | | : | : | | | | | | | | | |
Db 88 NELNCTKSVSPMEDKVMSCCPKDWRLFGSHCYLVPTVSSASWKNSESCRMGAHLVVI 147

Qy 118 NTEAEQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVRFVHLGEPNHSAQCA 177
      : : : | | | | : | : | | | | : | | | | | | | | | | | |
Db 148 QSOEQDITGILDTHAAYFGLMD-TGHRQWQVDQTPYEESTFVHNGEPSSGNEKCA 206

Qy 178 SIVF-WKPTGWNVDVICETRRNSICEMNKIYL 209
      : | : | | | | | | : | : | : | | | | | | | | | |
Db 207 TIIYRWK-TGWNWDISCLQKSKVCOMKINL 238

RESULT 13
Q9H229
ID Q9H229 PRELIMINARY; PRT; 204 AA.
AC Q9H229;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-type lectin DBB27 short form.
GN DBB27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN RP
RP SEQUENCE FROM N.A.
RA Richard M., Beaulieu A.D.;
RT "A short ORF for the C-type lectin DBB27.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200738; AAG35593.1; -.
DR InterPro; IPR002353; AntifreezeZell.
DR Pfam; PF00059; Lectin_C.1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS00615; C_TYPE_LECTIN.1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 204 AA; 23550 MW; 4DFEC27549CEAA5 CRC64;

Query Match 37.2%; Score 433.5; DB 4; Length 204;
Best Local Similarity 53.7%; Pred. No. 3.3e-36;
Matches 73; Conservative 22; Mismatches 40; Indels 1; Gaps 1;

Qy 75 AWGCCPASWKSFGSCYFISSEKVKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESF 134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 AWSCCPKNWKFSNCFISTESASWQSEKDCARMEAHLLVINTQEQDFIQNLQES 128

Qy 135 SYFLGLSDPOGNNWQIDKTPYEKNVRFVHLGEPNHSAQCAIVFWK-PTGWNWDVI 193
      : | | | | | | : | | | | | | : | | | | | | : | | | | | |
Db 129 AYFVGLSDPEGQHWQVDQTPYNESTFWHPREPSDPNERCVVLFNFKSPKRWGNDVN 188

Qy 194 CETRRNSICEMNKIYL 209
      | : | : | | | | | |
Db 189 CLGQSVCEMMKIHLL 204

RESULT 14
Q923C7
ID Q923C7 PRELIMINARY; PRT; 262 AA.
AC Q923C7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to C-type (calcium dependent, carbohydrate recognition
DE domain) lectin, superfamily member 6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN RN
RP SEQUENCE FROM N.A.
```

```

RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006623; AAH06623.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN.1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 262 AA; 29822 MW; 413DAD69511071B8 CRC64;

Query Match 36.9%; Score 430; DB 11; Length 262;
Best Local Similarity 37.2%; Pred. No. 1e-35;
Matches 89; Conservative 40; Mismatches 72; Indels 38; Gaps 8;

Qy 3 QEOQPOSTEKRGWLSURLMSVAGISTALLSACFIVSCVVTYHFTYGE----- 49
      : | : | | | | : | | | | | : | | | | | : | | | | | :
Db 30 REAPIRDLRKPGPSULLTSLM-LLLLLLAIFLVAFII-YFOKYSQLLEEKKAANNIMH 87

Qy 50 -----TGKRLSELHSYHSLTCTFSSEGTVPWAWGCCPASWKSFGSCYFI 93
      : | | | | | | : | | | | | : | | | | | : | | | | | :
Db 88 NELNCTKSVSPMEAPPFIGORALTLESIEDLGLAPEDKV--WSCCPKDWRLFGSHCYLV 145

Qy 94 S---SEKVKSKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESFSYFLGLSDPOGNNWQ 151
      | : | | | | | | | | | | : | | | | : | | | | | : | | | |
Db 146 PTVSSASWKNSESCRMGAHLVVIQSOEQDITGILDTHAAYFGLMD-TGHRQWQ 204

Qy 152 IDKTPYEKNVRFVHLGEPNHSAQCAIVF-WKPTGWNWDVICETRRNSICEMNKIYL 209
      : | | | | | | | | | | : | | | | | : | | | | | : | | | |
Db 205 VDOTPYEESITFWHNGEPSSGNEKCATIIRWK-TGWNWDISCLQKSKVCOMKINL 262

RESULT 15
Q8WXW9
ID Q8WXW9 PRELIMINARY; PRT; 230 AA.
AC Q8WXW9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fc-epsilon receptor III.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN RP
RP SEQUENCE FROM N.A.
RA Zhang W., Zhang J., Li N., Wan T., Chen T., Cao X.;
RT "Novel human Fc-epsilon receptor III.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF328684; AAL56016.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 230 AA; 26603 MW; 03CAB887D71BB084 CRC64;

Query Match 36.5%; Score 425; DB 4; Length 230;
Best Local Similarity 44.2%; Pred. No. 2.8e-35;
Matches 76; Conservative 31; Mismatches 61; Indels 4; Gaps 3;

Qy 26 ISIALLSACFIVSCVVTYHFTYGETGKRLSELHSYHSLTCTFSEGTKV--PAMGCCPASW 83
      | | | | | | : | | | | | : | | | | | : | | | | | :
Db 52 IFFLLAISFFIAFVIFQ-KYSQLEKTKTKELVHTTLECVKKNPVETAMSCCPKNW 110

Qy 84 KSGSCYFISSEKVKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESFSYFLGLSDP 143
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 111 KFSNCSYFISTESASWQSEKDCARMEAHLLVINTQEQDFIQNLQESAYFVGLSDP 170

Qy 144 OGNNNWQIDKTPYEKNVRFVHLGEPNHSAQCAIVFWK-PTGWNWDVIC 194
      : | : | | | | | | : | | | | | : | | | | | : | | | | | :
Db 171 EGQRHWQWVDQTPYNESTFWHPREPSDPNERCVVLFNFKSPKRWGNDVNC 222

Search completed: February 11, 2003, 23:00:24
```





```

RN  [1]
RC  SEQUENCE FROM N.A.
RP  STRAIN=BALB/C;
RA  Arizumi K., Shen G., Shikano S., Ritter R. III, Zukas P.,
RT  Edelbaum D., Morita A., Takashima A.;
RT  "Cloning of a second dendritic cell-associated C-type lectin (dectin-
RL  2) and its alternatively spliced isoforms.";
RL  J. Biol. Chem. 0:0-0(2000).
DR  EMBL: AF240359; AAF67179.1; -.
DR  HSSP: P23806; 11XX.
DR  InterPro: IPR001304; Lectin_C.
DR  Pfam: PF00059; lectin_c; 1.
DR  SMART: SM00034; CLECT; 1.
DR  PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
SQ  SEQUENCE 168 AA; 19332 MW; CF2ADDFE054EB038 CRC64;

Query Match      50.9%; Score 592.5; DB 11; Length 168;
Best Local Similarity 55.4%; Pred. No. 1.6e-52;
Matches 118; Conservative 16; Mismatches 30; Indels 49; Gaps 5;

QY  1 MMQEQQPOSTEKRG--WLSRLWSVAGISIALLSACPIVSCVVTYHFTYGETGKRLSELH 58
DB  1 MVQERQSGQ---KGVCW-TLRWGAAVISMLLTCTFASCVVITYQFIMQDPSRLLYELH 56
QY  59 SYHSSLTCFSEGTKV--PAWGCCPASWKSFGSSCYFISSEKVKWSSEQNCVEMGAHLV 116
DB  57 TYHSSLTCFSEGTWSEKWKMGCCPNHWKSKFGSSCYLISTRENFSTSEQNCVQGAHLV 116
QY  117 FNTAEQNFTVQOLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQC 176
DB  117 INTEAEQNFTVQOLNESLSYFLGLSN----- 142
QY  177 ASIVFMPKPTGWNVDVICETRRNSICEMNKIYL 209
DB  143 -----PSKMGWNDVFCDSKHNSICEMKKIYL 168

RESULT 5
Q8WXX8
ID  Q8WXX8 PRELIMINARY; PRT; 182 AA.
AC  Q8WXX8;
DT  01-MAR-2002 (TrEMBLrel. 20, Created)
DT  01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Dendritic lectin b isoform.
GN  CLEC5P11.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21426806; PubMed=11536172;
RA  Arce I., Roda-Navarro P., Montoya M.C., Hernanz-Falcon P.,
RA  Puig-Kroger A., Fernandez-Ruiz E.;
RA  "Molecular and genomic characterization of human DLEC, a novel member
RT  of the C-type lectin receptor gene family preferentially expressed on
RT  monocyte-derived dendritic cells.";
RL  Eur. J. Immunol. 31:2733-2740(2001).
DR  EMBL: AF325460; AAL37359.1; -.
DR  InterPro: IPR001304; Lectin_C.
DR  Pfam: PF00059; lectin_c; 1.
DR  SMART: SM00034; CLECT; 1.
DR  PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR  PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
KW  Lectin.
SQ  SEQUENCE 182 AA; 21498 MW; 78EB4E4E0109AE7 CRC64;

Query Match      45.9%; Score 534.5; DB 4; Length 182;
Best Local Similarity 54.7%; Pred. No. 1.4e-46;
Matches 94; Conservative 23; Mismatches 52; Indels 3; Gaps 1;

QY  41 VTYHFTYGETGKRUSEL---HSYHSSLTCFSEGTKVPWAGCCPASWKSFGSSCYFISSE 97
```

```

DB  11 VPHNFMYSKTVRLSKLREYQOYHPSLTCVMEGKDIEDMSCCPTPTWTSQSSCYFISTGM 70
QY  98 KYWKSKEQNCVEMGAHLVVFNTAEQNEFTVQOLNESFSYFLGLSDPGQNNWQWIDKTPY 157
DB  71 QSMTKSOKNCVSGADLVVINTREEQDFIQLNKRNSYFLGLSDPGGRRHWQWVDOTPY 130
QY  158 EKNVRFWHLGEPNHSAEQACVTFWKPRTGWNVDVICETRRNSICEMNKIYL 209
DB  131 NENVTFHSGEPNNLDERCAIINFRSSEKGNWDIHCYVPQKSICKMKKIYI 182

RESULT 6
Q9D8Q7
ID  Q9D8Q7 PRELIMINARY; PRT; 176 AA.
AC  Q9D8Q7;
DT  01-JUN-2001 (TrEMBLrel. 17, Created)
DT  01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  1810046I24Rik protein.
GN  1810046I24RIK.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX  MEDLINE=21085960; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA  Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA  Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA  Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA  Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA  Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA  Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA  Hayashizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  EMBL: AK007794; BAB25260.1; -.
DR  HSSP: P05140; 2AFP.
DR  MGD: MGI:1917060; 1810046I24Rik.
DR  InterPro: IPR001304; Lectin_C.
DR  Pfam: PF00059; lectin_c; 1.
DR  SMART: SM00034; CLECT; 1.
DR  PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR  PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
SQ  SEQUENCE 176 AA; 20342 MW; C020F749633D27BA CRC64;

Query Match      40.8%; Score 475.5; DB 11; Length 176;
Best Local Similarity 43.6%; Pred. No. 1.4e-40;
Matches 92; Conservative 30; Mismatches 52; Indels 37; Gaps 6;

QY  1 MMQEQQPOSTEKRGWLSRLWSVAGISIALLSACPIVSCVVTYHFTYGETGKRLSELH 60
DB  1 MVQERLQG-KAVSW-SLRWSAAVISILLSTCFIASCV----- 38
QY  61 HSSLTCFSEGTKVPWAGCCPASWKSFGSSCYFISSEKVKWSKQNCVEMGAHLVFN 118
DB  39 -----DKV--WSCCPKDWKLFSGHCYLVPTVFSSASWNKSEENSRMGHLVVIH 86
QY  119 TEAEQNFTVQOLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQAC 178
DB  119 TEAEQNFTVQOLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQAC 178
```

```
Best Local Similarity 68.1%; Pred. No. 7.6e-73;
Matches 145; Conservative 19; Mismatches 41; Indels 8; Gaps 4;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFTVSCVVTYHFTYGETGKRLSELH 58
Db 1 MVQERQSG---KGVCW-TLRLWSAAVISMLLSTCFIASCVVYQFIMDQPSRRILYELH 56

Qy 59 SYHSSLTCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 116
Db 57 TYHSSLTCFSEGTGVSKWGCPCPNHWKSGSGCYLSTKENFWSVSEQNCVQMGAGHLVV 116

Qy 117 FNTAEQNFIVQOLNESFSYFLGLSDPQGNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 176
Db 117 INTEAEQNFITQOLNESLSYFLGLSDPQGNKQWIDDTFSONVRFWHPHPNLPPEERC 176

Qy 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 177 VSIYVWNPCKWGNVDVFCDSKHNSECIMCKKIYL 209

RESULT 2
Q9JKF3 PRELIMINARY; PRT; 175 AA.
AC Q9JKF3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Dectin-2 beta isoform.
GN CLCsf10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Arizumi K., Shen G., Shikano S., Ritter R. III, Zukas P.,
Edelbaum D., Morita A., Takashima A.;
RT "Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms.";
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AF240358; AAF67178.1; -.
DR HSSP; P22897; IEGG.
DR MGI; 1861231; Clcfsf10.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 175 AA; 20313 MW; EIC9632687BF0EA9 CRC64;

Query Match 56.6%; Score 659; DB 11; Length 175;
Best Local Similarity 58.3%; Pred. No. 2.8e-59;
Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFTVSCVVTYHFTYGETGKRLSELH 58
Db 1 MVQERQSG---KGVCW-TLRLWSAAVISMLLSTCFIASCV----- 38

Qy 59 SYHSSLTCFSEGTGVKVPANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVVFN 118
Db 39 -----EKNWGCPCPNHWKSGSGCYLSTKENFWSVSEQNCVQMGAGHLVVFN 84

Qy 119 TEAEQNFIVQOLNESFSYFLGLSDPQGNMOWIDKTPYEKNVRFWHLGEPNHSAEQAS 178
Db 85 TEAEQNFITQOLNESLSYFLGLSDPQGNKQWIDDTFSONVRFWHPHPNLPPEERCVS 144

Qy 179 IVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 145 IVYVWNPCKWGNVDVFCDSKHNSECIMCKKIYL 175

RESULT 3
Q9JKF2 PRELIMINARY; PRT; 168 AA.
AC Q9JKF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dectin-2 gamma isoform.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

Q8WTT0 PRELIMINARY; PRT; 213 AA.
AC Q8WTT0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Dendritic lectin (Blood dendritic cell antigen 2 protein).
GN CLCsf11 OR BDCA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21426806; PubMed=11536172;
RA Arce I., Roda-Navarro P., Montoya M.C., Hernanz-Falcon P.,
Puig-Kroger A., Fernandez-Ruiz E.;
RT "Molecular and genomic characterization of human DLEC, a novel member of the C-type lectin receptor gene family preferentially expressed on monocyte-derived dendritic cells.";
RT Eur. J. Immunol. 31:2733-2740(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Sohma Y., Johnston I., Dzionek A., Nagasaka T., Nagafune J., Okada T.,
Hirano T., Schmitz J., Yamaguchi Y.;
RT "Cloning and characterization of blood dendritic cell antigen 2 (BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF325459; AAL37358.1; -.
DR EMBL; AF293615; AAL37036.1; -.
DR InterPro; IPR003654; GASA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF02704; GASA; 1.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 213 AA; 25037 MW; 5DC82C95BE2378C1 CRC64;

Query Match 54.7%; Score 637; DB 4; Length 213;
Best Local Similarity 54.0%; Pred. No. 6.2e-57;
Matches 115; Conservative 30; Mismatches 64; Indels 4; Gaps 2;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFTVSCVVTYHFTYGETGKRLSEL-- 57
Db 1 MVPEEPQDREKGLWFWLQKWSMAVVSILLSVCTVSSVVPHPNPMYKTVKRLSKLRE 60

Qy 58 -HSYHSSLTCFSEGTGVKVPANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 116
Db 61 YQYHFSLTCVMEGKDIEDWSCCPTPTSFQSSCYFISTGMOSWTQSKNCSVMGADLVV 120

Qy 117 FNTAEQNFIVQOLNESFSYFLGLSDPQGNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 176
Db 121 INTREQDFIQLKNSSYFLGLSDPQGRHQQWVDQTPYENVTFWHSGEPNLDERC 180

Qy 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 181 AINFRSSEWGNIDHCHVPQKICKMKKIYI 213

RESULT 4
Q9JKF2 PRELIMINARY; PRT; 168 AA.
AC Q9JKF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dectin-2 gamma isoform.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:55:56 : Search time 30 seconds  
(without alignments)  
1435.462 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRGWLSIRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-rvirus:\*  
16: sp-bacteriap:\*  
17: sp-archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	793	68.1	209	11	Q9JKF4	Q9jkf4 mus musculus
2	659	56.6	175	11	Q9JKF3	Q9jkf3 mus musculus
3	637	54.7	213	4	Q8WTF0	Q8wtf0 homo sapien
4	592.5	50.9	168	11	Q9JKF2	Q9jkf2 mus musculus
5	534.5	45.9	182	4	Q8WXX8	Q8wxw8 homo sapien
6	475.5	40.8	176	11	Q9D8Q7	Q9d8q7 mus musculus
7	458	39.3	237	4	Q9UT34	Q9uli34 homo sapien
8	458	39.3	237	4	Q9UMR7	Q9umr7 homo sapien
9	453	38.9	237	4	Q9NS33	Q9ns33 homo sapien
10	443	38.0	219	11	Q9Z2H6	Q9z2h6 mus musculus
11	441	37.9	215	4	Q8WX18	Q8wx18 homo sapien
12	436	37.4	238	11	Q9QZ15	Q9qz15 mus musculus
13	433.5	37.2	204	4	Q9H229	Q9h2z9 homo sapien
14	430	36.9	262	11	Q9Z3C7	Q9z3c7 mus musculus
15	425	36.5	230	4	Q8WXX9	Q8wxw9 homo sapien
16	398.5	34.2	219	4	Q9ULY5	Q9uly5 homo sapien

## ALIGNMENTS

RESULT 1

Q9JKF4 PRELIMINARY; PRT; 209 AA.  
ID Q9JKF4  
AC Q9JKF4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Dectin-2 alpha isoform (C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10).  
GN CLEGSF10.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C.  
RA Arizumi K., Shen G., Shikano S., Ritter R. III, Zukas P., Edelbaum D., Morita A., Takashima A.;  
RT "Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms.";  
RL J. Biol. Chem. 0:0-0(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF240357; AAF67177.1; -.  
DR EMBL; BC023008; AAH23008.1; -.  
DR HSSP; P22897; 1EGG.  
DR MGD; MGI:1861231; Clecsf10.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; UNKNOWN\_1.  
DR PROSITE; PS50041; C\_TYPE\_LLECTIN\_2; 1.  
KW Lectin.  
SQ SEQUENCE 209 AA; 24324 MW; 7AB8BDC0DB1EA5ED CRC64;

Query Match 68.1%; Score 793; DB 11; Length 209;



Qy 165 HLGEPNH---SAEQCASIVFWKPTGNGWNDVICETRNSICEMNKI 207  
 : : : : : : : : : : : : : : : : : : :  
 Db 771 NPGCPDPSYLSGENCV-VVMVHHQQ-QWSDVPNCYHLSTCKMGLV 814

Search completed: February 11, 2003, 23:01:57  
Job time : 30 secs

```
FT VARSPLIC 646 883 MISSING (IN GPI-ANCHORED ISOFORM).
FT CONFLICT 51 52 AL-> WV (IN REF. 3).
FT CONFLICT 503 503 V-> L (IN REF. 2).
FT CONFLICT 519 509 TV-> PA (IN REF. 2).
FT CONFLICT 526 526 G-> R (IN REF. 2).
FT CONFLICT 541 541 G-> A (IN REF. 2).
FT CONFLICT 556 556 R-> S (IN REF. 2).
FT CONFLICT 573 573 E-> A (IN REF. 2).
FT CONFLICT 583 583 V-> L (IN REF. 2).
FT CONFLICT 649 649 V-> L (IN REF. 2).
FT CONFLICT 670 670 P-> A (IN REF. 2).
FT CONFLICT 738 738 P-> A (IN REF. 2).
FT CONFLICT 809 809 R-> A (IN REF. 2).
SQ SEQUENCE 883 AA; 96057 MW; AC7ACC40CB53ED37 CRC64;

Query Match 15.7%; Score 183; DB 1; Length 883;
Best Local Similarity 26.1%; Pred. No. 7e-09;
Matches 62; Conservative 31; Mismatches 83; Indels 62; Gaps 13;

Qy 3 QEQQPQSTKRGWLSRLWSV-----AGISIALLSA-CFIVSCVVYHFTYG 48
Db 581 REVPTEPKSGRVLGTGSVQAPVLPTDSASRGVAVAPSSGDCIPSPC----- 631
Qy 49 ETGRKLSLHSHSLTCTF--SEGTG---VPAWG---C-----CPASWKSPGSSCYFIS 94
Db 632 -----HNGGTCLEBEKEGFRCLCVPGYGGDLCDVGLHFCSPGWEPFGACVYKHF 679
Qy 95 SEEKVSKSEONCVEMGAHLVVFTEAEONFIVQOLNESFSYFLGLSDPGNNWQWIDK 154
Db 680 STRSRWEAESQCRALGALHTSICTPEQDFVNDYRE--YQWIGLNDRTIEGDFLWSGD 737
Qy 155 TP--YEKNVRVHLGEPNH-----SAEQASIVFWKPTGWNQDVICETRRNSICEMNKI 207
Db 738 PPLLYEN----WNPGQPDYSFLSGENCV-VMWIHDQG-QMSDVPNCNHLSTCKMGLV 789

RESULT 15
PCGB_BOVIN STANDARD; PRT; 912 AA.
AC Q28062;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brevican core protein precursor.
CN BCAN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94193597; PubMed=8144512;
RA Yamada H., Watanabe K., Shimonaka M., Yamaguchi Y.;
RT "Molecular cloning of brevicin, a novel brain proteoglycan of the
RT agrecan/versican family.";
RL J. Biol. Chem. 269:10119-10126(1994).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- TISSUE SPECIFICITY: BRAIN; EXPRESSED IN CEREBELLAR ASTROCYTES BUT
CC NOT IN NEURONS.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X75887; CAA53481.1; -.
CC HSSP; P20693; 1HLJ.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_V.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00406; IGV; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC PROSITE; PS01241; LINK; 2.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 912 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 32 158 LINK 1.
FT DOMAIN 174 251 LINK 2.
FT DOMAIN 272 353 EGF-LIKE.
FT DOMAIN 647 683 C-TYPE LECTIN.
FT DOMAIN 683 811 SUSHI.
FT DOMAIN 812 876 BY SIMILARITY.
FT DISULFID 57 137 BY SIMILARITY.
FT DISULFID 179 250 BY SIMILARITY.
FT DISULFID 203 224 BY SIMILARITY.
FT DISULFID 277 352 BY SIMILARITY.
FT DISULFID 301 322 BY SIMILARITY.
FT DISULFID 651 662 BY SIMILARITY.
FT DISULFID 656 671 BY SIMILARITY.
FT DISULFID 673 682 BY SIMILARITY.
FT DISULFID 689 700 BY SIMILARITY.
FT DISULFID 717 809 BY SIMILARITY.
FT DISULFID 785 801 BY SIMILARITY.
FT DISULFID 816 859 BY SIMILARITY.
FT DISULFID 845 872 BY SIMILARITY.
FT CARBOHYD 130 130 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 912 AA; 99554 MW; 677B3EB1C688C4D7 CRC64;

Query Match 15.7%; Score 182.5; DB 1; Length 912;
Best Local Similarity 29.5%; Pred. No. 8.1e-09;
Matches 49; Conservative 26; Mismatches 64; Indels 27; Gaps 10;

Qy 61 HSSLTCE--SEGTG---VPAWG---C-----CPASWKSPGSSCYFISSEKWKSEQN 106
Db 657 HNGGTCLEBEKEGFRCLCVPGYGGDLCDVGLHFCSPGWDAFQGCYKHFARRSWEEAKN 716
Qy 107 CVMGAHLVVFTEAEONFIVQOLNESFSYFLGLSDPGNNWQWIDKTP--YEKNVRWF 164
Db 717 CRMYGAHLASISTPEQDFINNRYRE--YQWIGLNDRTIEGDFLWSGDVGPLLYEN---W 770
```

```
Query Match 15.7%; Score 183; DB 1; Length 883;
Best Local Similarity 25.6%; Pred. No. 7e-09;
Matches 61; Conservative 33; Mismatches 82; Indels 62; Gaps 13;

QY 3 QEQQSQTEKRGWLSRLWSV-----AGISIALLSA-CFIVSCVVTYHTYTG 48
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 581 RELETPSEKSGRTVLAGTSVQAQVLPVLTDSASHGVAVAPSSGDCIPSPC----- 631
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 49 ETGKRLSELHSHYSSLTGCF--SEGTK--VPAMG--C-----CPASWKSFGSSCYFIS 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 632 -----HNGGTGLEKEGFRCLCPGYGDLCDVGLHFCSPGWEAFQACVYKH 679
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 SEEKVWSKSEQNCVEMGAHLVVFVTEAEQNFIVQOLNSESFYGLSDPOCQNNWQWIDK 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 680 STRSWEAEASOCRALGAHLTSICTPEQDFVNDYRE--YOWIGLNDRTIEGDFLWSDG 737
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 TP--YEKNVRWHIGEPNH----SAEQCASIVFWKPTGMGNVDVTCETFRNSICBMNKI 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 738 APLLYEN----WNPQPDSPYLSGNCV-VNVVHDOG-QWSDVPCNVHLSYCKMGLV 789
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
PCGB_RAT STANDARD; PRT; 883 AA.
AC P55068; Q63040; Q62860; Q63513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brevican core protein precursor (Brain enriched hyaluronan binding protein) (BEHAV protein).
GN BCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN
SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUP=Brain;
RX MEDLINE=96070828; PubMed=7592978;
RA Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C., Gundelfinger E.D.;
RT "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as secreted and cell surface glycosylphosphatidylinositol-anchored isoforms.";
RL J. Biol. Chem. 270:27206-27212(1995).
RN [2]
SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RC TISSUP=Brain;
RX MEDLINE=96074575; PubMed=7488217;
RA Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamauchi Y.;
RT "cDNA cloning and the identification of an aggrecanase-like cleavage site in rat brevican.";
RL Biochem. Biophys. Res. Commun. 216:957-963(1995).
RN [3]
SEQUENCE OF 1-423 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUP=Brain;
RX MEDLINE=94216386; PubMed=7512973;
RA Jaworski D.M., Kelly G.M., Hockfield S.;
RT "BEHAV, a new member of the proteoglycan tandem repeat family of hyaluronan-binding proteins that is restricted to the brain.";
RL J. Cell Biol. 125:495-509(1994).
CC
CC -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN. THE GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-BEARING CELL SURFACE RECEPTOR.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A SECRETED FORM (SHOWN HERE) AND A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.
```



CC	Anthocidaris.
OX	NCBI_TaxID=7629;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=87194838; PubMed=3571253;
RA	Giga Y., Ikal A., Takahashi K.;
RT	"The complete amino acid sequence of echinoidin, a lectin from the
FT	coelomic fluid of the sea urchin <i>Anthocidaris crassispina</i> . Homologies
FT	with mammalian and insect lectins.";
RL	J. Biol. Chem. 262:6197-6203(1987).
CC	-!- FUNCTION: ROLE IN THE DEFENCE SYSTEM OF THE ORGANISM AGAINST
CC	MICROORGANISMS. THIS LECTIN IS SPECIFIC TO GAL-GAL(NAC).
CC	-!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC	-!- TISSUE SPECIFICITY: COELOMIC FLUID.
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
DR	HSP; A26697; A26697.
DR	HSP; P20693; 1HLJ.
DR	InterPro: IPR001304; Lectin_C.
DR	Pfam: PF00059; lectin_c; 1.
DR	SMART: SM00034; CLECT; 1.
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
KW	Glycoprotein; Lectin.
FT	DOMAIN 1 143 C-TYPE LECTIN (LONG FORM).
FT	DISULFID 2 2 INTERCHAIN.
FT	DISULFID 3 14
FT	DISULFID 31 141
FT	DISULFID 116 132
FT	CARBOHYD 38 38
FT	SITE 39 41 O-LINKED.
FT	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE 147 AA; 16661 MW; 849E98645481A062 CRC64;
Query Match 16.7%; Score 194.5; DB 1; Length 147;	
Best Local Similarity 31.3%; Pred. No. 8.3e-11;	
Matches 46; Conservative 56; Mismatches 58; Indels 23; Gaps	
Qy	77 GCCPASWKSFGSSCYFTSSBEKVSKSEQNC-----VEMGAHLVVFENTEAEQNF 120
Db	
Qy	127 VQQLNESF-----SYFLGLSDPGNNWMIKTPYKKNVRFWHLGPNH--SAEOCAS 178
Db	
Qy	179 IVFKWPTGWNDVIC-ETRNSICEM 204
Db	
Db	119 MVM--GAGLNWIDLPCSTRHYLCKL 143
RESULT 12	
KUCR	MOUSE
ID	KUCR_MOUSE STANDARD; PRT; 548 AA.
AC	P70194;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	C-type lectin I3 (Kupffer cell receptor).
GN	CLECSP13 OR KCLR.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BALB/c; TISSUE=Liver;
RA	Takezawa R., Wagatsuma H., Nomoto C., Watanabe Y., Akaike T.;
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
CC	COULD BE INVOLVED IN ENOCYTOSIS.
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.
CC	-!- TISSUE SPECIFICITY: KUPFFER CELLS.
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN

Db 260 KQPDNWNHGLGGEDCAHFT----SDGRWDDVQCRPYRWCEM 300

# RESULT 9

```

ID  LECH_MOUSE  STANDARD;  PRT;  283 AA.
AC  P34927; Q64363;
DT  01-FEB-1994 (Rel. 28, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Asialoglycoprotein receptor 1 (Hepatic lectin 1) (MHL-1) (ASGP-R)
DE  (ASGPR).
GN  ASGR1 OR ASGR-1.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93176818; PubMed=8439566;
RA  Takezawa R., Shirazawa K., Watanabe Y., Akaike T.;
RT  "Determination of mouse major asialoglycoprotein receptor cDNA
RT  sequence.";
RL  Biochim. Biophys. Acta 1172:220-222(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=BALB/C; TISSUE=Liver;
RC  MEDLINE=95047431; PubMed=7958950;
RA  Monroe R.S., Huber B.E.;
RT  "The major form of the murine asialoglycoprotein receptor: cDNA
RT  sequence and expression in liver, testis and epididymis.";
RL  Gene 148:237-244(1994).
CC  -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC  WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC  CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC  TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE UNITS. AFTER LIGAND
CC  BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC  TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC  DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC  SURFACE.
CC  -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC  -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC  CELLS.
CC  -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC  -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D13517; BAA02734.1; -
DR  EMBL; U09362; AAB60441.1; -
DR  EMBL; U08372; AAB60440.1; -
DR  PIR; S29855; S29855.
DR  HSP; P20693; 1HLJ.
DR  MGD; MGI:88081; Asgr1.
DR  InterPro; IPR001304; Lectin_C.
DR  Pfam; PF00059; lectin_c; 1.
DR  SMART; SM00034; CLECT; 1.
DR  PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR  PROSITE; PS50041; C-TYPE LECTIN 2; 1.
KW  Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW  Calcium; Signal-anchor; Phosphorylation.
FT  INIT MET 0 0 BY SIMILARITY.
FT  DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 39 59 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT  (POTENTIAL).
FT  DOMAIN 60 283
FT  DOMAIN 151 277 C-TYPE LECTIN (LONG FORM).

```

```

FT SITE 4 7 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 152 163 BY SIMILARITY.
FT DISULFID 180 275 BY SIMILARITY.
FT DISULFID 253 267 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 150 150 I -> T (IN REF. 1).
SQ SEQUENCE 283 AA; 32472 MW; 982A5D305AAE0D8F CRC64;

Query Match 18.6%; Score 217; DB 1; Length 283;
Best Local Similarity 31.6%; Pred. No. 1.6e-12;
Matches 42; Conservative 25; Mismatches 52; Indels 14; Gaps 4;

```

```

QY 78 CCPASWKSPGSSCYFISSEKVKSKSFONCVEMGAHLVVFENTEAEQNFIVQQLNESFSYF 137
Db 151 CCPINWVEYEGSCYWFSSSRPMTADKYCOLENAHLVVVTSRDEQNFQRMHG-PLNTW 209
QY 138 LGLSDPOGNNWOWIDKTPYEKNVRFWHLGEPNH-----SAEQCAISIVFWKPTGWN 190
Db 210 IGLTDQ--NGPWKWDGTDYETGFQNRPPQPDNWNVYHGLGGEDCAHFT----TDGRWN 263
QY 191 DVICETFRNSICE 203
Db 264 DDVCRPRYRWCE 276

```

## RESULT 10

```

ID  LECH_HUMAN  STANDARD;  PRT;  290 AA.
AC  P07306;
DT  01-APR-1988 (Rel. 07, Created)
DT  01-APR-1988 (Rel. 07, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Asialoglycoprotein receptor 1 (Hepatic lectin H1) (ASGPR) (ASGP-R).
GN  ASGR1.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=85130911; PubMed=2982798;
RA  Spiess M., Schwartz A.L., Lodish H.F.;
RT  "Sequence of human asialoglycoprotein receptor cDNA. An internal
RT  signal sequence for membrane insertion.";
RL  J. Biol. Chem. 260:1979-1982(1985).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86079574; PubMed=3753585;
RA  Spiess M., Lodish F.;
RT  "An internal signal sequence: the asialoglycoprotein receptor membrane
RT  anchor.";
RL  Cell 44:177-185(1986).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  Wang H., Gao X., Li L., Lou H., Huang Y., Wang B., Han J.;
RT  "Human asialoglycoprotein receptor 1 gene is expressed in SH-SY5Y
RT  human neuroblastoma cells.";
RL  Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC  WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC  CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC  TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE UNITS. AFTER LIGAND
CC  BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC  TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC  DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC  SURFACE.
CC  -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC  -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC  CELLS.
CC  -!- PTM: PHOSPHORYLATED ON A CYTOPLASMIC SER RESIDUE.
CC  -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.

```

```
CC CC SURFACE.
CC CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CC CELLS.
CC CC -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC CC -!- MISCELLANEOUS: TWO TYPES OF RAT HEPATIC LECTIN HAVE BEEN
CC CC IDENTIFIED, RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.
CC CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: K02817; AAA42037.1; -.
CC CC EMBL: M21770; AAA40764.1; -.
CC CC PIR: A03166; LNRTL.
CC CC HSP: P20693; IHLJ.
CC CC InterPro: IPR001304; Lectin_C.
CC CC Pfam: PF00059; lectin_c; 1.
CC CC SMART: SM00034; CLECT; 1.
CC CC PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
CC CC PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
CC CC LECTIN; Glycoprotein; Receptor; Endocytosis; Transmembrane;
CC CC Calcium; Signal-anchor; Phosphorylation.
CC CC KW
CC CC INIT MET 0
CC CC DOMAIN 1 38
CC CC TRANSMEM 39 59
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC EXTRACELLULAR (POTENTIAL).
CC CC C-TYPE LECTIN (LONG FORM).
CC CC ENDOCYTOSIS SIGNAL (POTENTIAL).
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC O -> R (IN REF. 2).
CC CC CONFLICT 60 60
CC CC SEQUENCE 283 AA; 32718 MW; 3BA2631A5E28A993 CRC64;
CC CC
Query Match 19.4%; Score 226.5; DB 1; Length 283;
Best Local Similarity 25.9%; Pred. No. 2.3e-13;
Matches 64; Conservative 31; Mismatches 69; Indels 83; Gaps 11;
QY 25 GISALLSACFIVSCVTVTH-----FTYGE-TGKRL 54
DB 45 GLSILL-----VVVCVITSONSLRDLRLVRQNFSTVSTEDQVKALTTQGERVGRKM 100
QY 55 -----SELHSYHS-----SLTCFSEGGKVPAMG-----CCPASW 83
DB 101 KLIVESOLEKHQEDLRDREHLLHVKQLVSDVRSLSQMAALR-----GNGSRICCPINW 156
QY 84 KSFSGSYFISBEKVKWSEQNCVEMGAHLVVFNTAEQNFIVQQLNESFYFLGLSDP 143
DB 157 VEYEGSGYFWSSSVKPTWADKYQLENAHLVVVTSWEQRFVQOHMG-PLNTWIGLTDQ 215
QY 144 QGNKNWQIDKTPYEKNVRFWHLGEPNH-----SAEQCASIVFWKPTGWNVDVICET 196
DB 216 --NGPKWVDGTDYETGFKNWRPGPDWYGHGLGGEDCAHFT----TDGHNDDVCR 269
QY 197 RRNSICE 203
DB 270 PYRWVCE 276
RESULT 8
MMGL_RAT
ID MMGL_RAT STANDARD; PRT; 306 AA.
AC P49301;
```

```
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
DE galactose/N-acetylgalactosamine-specific lectin) (MMGL).
GN MGL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90293078; PubMed=2358462;
RA Ii M., Kurata H., Itoh N., Yamashina I., Kawasaki T.;
RT "Molecular cloning and sequence analysis of cDNA encoding the
RT macrophage lectin specific for galactose and N-acetylgalactosamine.";
RL J. Biol. Chem. 265:11295-11298(1990).
RN [2]
RP PRELIMINARY SEQUENCE OF 9-28.
RX MEDLINE=88339956; PubMed=3421964;
RA Ii M., Kawasaki T., Yamashina I.;
RT "Structural similarity between the macrophage lectin specific for
RT galactose/N-acetylgalactosamine and the hepatic asialoglycoprotein
RT binding protein.";
RL Biochem. Biophys. Res. Commun. 155:720-725(1988).
CC -!- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE
CC UNITS.
CC -!- SUBUNIT: HOMO-OLIGOMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: J05495; AAA41216.1; -.
CC CC HSP: P06734; IHLI.
CC CC InterPro: IPR001304; Lectin_C.
CC CC Pfam: PF00059; lectin_c; 1.
CC CC SMART: SM00034; CLECT; 1.
CC CC PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
CC CC PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
CC CC LECTIN; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
CC CC DOMAIN 1 37
CC CC TRANSMEM 38 58
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC EXTRACELLULAR (POTENTIAL).
CC CC C-TYPE LECTIN (LONG FORM).
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC SEQUENCE 306 AA; 34242 MW; D68A5DFF0B9E8F13 CRC64;
CC CC
Query Match 19.1%; Score 222; DB 1; Length 306;
Best Local Similarity 31.5%; Pred. No. 6.4e-13;
Matches 52; Conservative 25; Mismatches 72; Indels 16; Gaps 5;
QY 49 ETGKRLSELHSYHSLSLTCFSEGTR--VPAMGCCPASWKSFGSSCYFISBEKVKWSEQN 106
DB 143 EITDRVQOLGKDLKTLTLCQLASLNKNGSAVACCPHLHMEHSGCYWFQSCKPMPADKY 202
QY 107 CVMGAHLVVENTAEQNFIVQQLNESFYFLGLSDPQGNNNWQIDKTPYEKNVRFWHL 166
DB 203 QLENSNLVVVNSLAEQNFILQTHMG-SVWVWIGLTDQ--NGPWRWVDGTDYKGFTHWAP 259
QY 167 GEPNH-----SAEQCASIVFWKPTGWNVDVICETRRNSICEM 204
DB 167 GEPNH-----SAEQCASIVFWKPTGWNVDVICETRRNSICEM 204
```

Db 505 NWRHNGEREDCVHLQRM-----WMDMACGTAYNWVCK 537

## RESULT 6

```

LECH_CHICK
ID LECH_CHICK STANDARD; PRT; 207 AA.
AC P02707;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hepatic lectin.
OS Gallus gallus (Chicken).
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RX MEDLINE=91268022; PubMed=2050668;
RA Bezouska K., Crichtow G.V., Rose J.M., Taylor M.E., Drickamer K.;
RT "Evolutionary conservation of intron position in a subfamily of genes
RT encoding carbohydrate-recognition domains."
RL J. Biol. Chem. 266:11604-11609(1991).
RN [2]
RX MEDLINE=88186849; PubMed=3281941;
RA Mellow T.E., Halberg D., Drickamer K.;
RT "Endocytosis of N-acetylglucosamine-containing glycoproteins by rat
RT fibroblasts expressing a single species of chicken liver glycoprotein
RT receptor."
RL J. Biol. Chem. 263:5468-5473(1988).
RN [3]
RX MEDLINE=81215504; PubMed=7240175;
RA Drickamer K.;
RT "Complete amino acid sequence of a membrane receptor for
RT glycoproteins. Sequence of the chicken hepatic lectin."
RL J. Biol. Chem. 256:5827-5839(1981).
CC -!- FUNCTION: HEPATIC LECTIN IS A MEMBRANE RECEPTOR PROTEIN THAT
CC RECOGNIZES AND BINDS EXPOSED N-ACETYLGLUCOSAMINE MOIETIES OF
CC PLASMA GLYCOPROTEINS, THUS MEDIATING THEIR CLEARANCE (FROM THE
CC CIRCULATION) AND ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- PTM: SOME OR ALL OF THE CYSTEINES ARE INVOLVED IN DISULFIDE BONDS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; M63230; AAA48814.1;
DR EMBL; M63225; AAA48814.1; JOINED.
DR EMBL; M63226; AAA48814.1; JOINED.
DR EMBL; M63227; AAA48814.1; JOINED.
DR EMBL; M63228; AAA48814.1; JOINED.
DR EMBL; M63229; AAA48814.1; JOINED.
DR EMBL; J03188; AAA48937.1;
DR PIR; A03167; LNCHL.
DR HSP; P20693; LHLJ.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN.1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN.2; 1.
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Signal-anchor; Acetylation.
FT MOD_RES 1 1 ACETYLATED.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 24 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT DOMAIN 49 207 (POTENTIAL).
FT FT DOMAIN 77 203 EXTRACELLULAR (POTENTIAL).
FT FT DISULFID 78 92 C-TYPE LECTIN (LONG FORM).
FT FT DISULFID 109 201 BY SIMILARITY.
FT FT DISULFID 179 193 BY SIMILARITY.
FT FT CARBOHYD 67 67 N-LINKED (GLCNAC...).
SQ SEQUENCE 207 AA; 24216 MW; 1F6B36FDB32895DB CRC64;

Query Match 19.8%; Score 231; DB 1; Length 207;
Best Local Similarity 29.6%; Pred. No. 6.3e-14;
Matches 56; Conservative 29; Mismatches 78; Indels 26; Gaps 7;

QY 18 LRLWSVAGISIALLSACFIYSCVYHYFTYGTCKRLKRLSELHSYHSSITCFSEGTQKYPAWG 77
DB 37 LLTLTLLSSVLARIAA-----LSSKLSTLQSEPKHNFSSRDSLLFP-C-G 79
QY 78 CCPASWKSFGSCYFISSEKQVMSKQNCVEMGAHLVAVFTEAEQNFIV-QOLNESFSY 136
DB 80 AQSQWEYFEGRCYFYFSLSRMSWHKAKAECEMHSHLIIDSYAKQNFVYFRTNERF-- 137
QY 137 FLGLSDPQGNNNQWIDKTPYKXNVRFWHLGEPNHS--EQCASIVFWKPTGWGNDVIC 194
DB 138 WIGLTDENQEGEWQWVDGTDTRSSFTFWKGEPPNRCFNEDCAHV--W--TSGWQNDVYC 193
QY 195 ETRNSICE 203
DB 194 TYECYVYCE 202

RESULT 7
LECH_RAT
ID LECH_RAT STANDARD; PRT; 283 AA.
AC P02706;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (RHL-1) (ASGP-R)
DE (ASGPR).
GN ASGR1 OR ASGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RX MEDLINE=86008335; PubMed=2995379;
RA Leung J.O., Holland E.C., Drickamer K.;
RT "Characterization of the gene encoding the major rat liver
RT asialoglycoprotein receptor."
RL J. Biol. Chem. 260:12523-12527(1985).
RN [2]
RX MEDLINE=85063786; PubMed=6095287;
RA Holland E.C., Leung J.O., Drickamer K.;
RT "Rat liver asialoglycoprotein receptor lacks a cleavable NH2-terminal
RT signal sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 81:7338-7342(1984).
RN [3]
RX MEDLINE=87026895; PubMed=2945599;
RA Watts C.;
RT "Isolation and expression of cDNA clones for a rat liver
RT asialoglycoprotein receptor."
RL Biosci. Rep. 6:527-534(1986).
CC -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE

```



```

RN [2]
RP SEQUENCE OF 102-120 AND 137-151.
RC STRAIN-C3H/HeN;
RX MEDLINE=89197865; PubMed=3241002;
RA Oda S., Sato M., Toyoshima S., Osawa T.;
RT "Purification and characterization of a lectin-like molecule specific
RT for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";
RL J. Biochem. 104:600-605(1988).
CC -!- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE
CC UNITS MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL
CC MACROPHAGES AND TUMOR CELLS.
CC
CC -!- SUBUNIT: HOMO-OLIGOMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED
CC MACROPHAGES.
CC
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S36676; AAB22171.1; -.
DR HSSP; P06734; 1HLI.
DR MGD; MGI:96975; Mgl.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
KW DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT FT
FT DOMAIN 57 304 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 172 298 C-TYPE LECTIN (LONG FORM).
FT DISULFID 173 184 BY SIMILARITY.
FT DISULFID 201 296 BY SIMILARITY.
FT DISULFID 274 288 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 304 AA; 34596 MW; 3F79CD12C34F5BCC CRC64;

Query Match 20.9%; Score 244; DB 1; Length 304;
Best Local Similarity 33.7%; Pred. No. 6.6e-15;
Matches 55; Conservative 27; Mismatches 59; Indels 22; Gaps 6;

Qy 54 LSELHSHSLTCTCFSEGTGVPAWGCCPASWKSFGSSCYFISSEKVKWSKSEQNCV 108
Db 146 VOOLRKDLKATCQLANLKNNGSEV---ACPLHWTEHSGCYWFSESEKSWPEADKYCR 202
Qy 109 EMGAHLVFEAEQNTIVQOLNESFSYFLGLSDPQGNNNQWIDKTP--YKKNVRFHILGE 168
Db 203 LENSILVNVNSLEEQNFQNLNANVSN-IGLTQ--NGPWRVVDGTDGFEKGFKNWAPLQ 259
Qy 169 PNH-----SAEQCASIVFWKPTGNGWINDVICETRRNSICE 204
Db 260 PDNWFHGLGGEDCAHIT----TGGPNDDVCQRTFRWICEM 298

RESULT 5
KUCR_RAT STANDARD; PRT; 550 AA.
AC P10716;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-type lectin 13 (Kupffer cell receptor).
GN CLECSF13 OR KCLR.
OS Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.
RX MEDLINE=88227939; PubMed=2836387;
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
RT binding receptor unique to rat Kupffer cells.";
RL J. Biol. Chem. 263:7487-7492(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91107689; PubMed=1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
RT rat Kupffer cells.";
RL J. Biol. Chem. 266:1850-1857(1991).
CC -!- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
CC COULD BE INVOLVED IN ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: KUPFFER CELLS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03734; AAA41472.1; -.
DR EMBL; M55532; AAA40892.1; -.
DR PIR; A28166; A28166.
DR PIR; A38674; A38674.
DR HSSP; P20693; 1HLU.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR00017; Syntaxin.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00503; SYN; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
KW Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 70 550 EXTRACELLULAR (POTENTIAL).
FT DISULFID 438 538 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;

Query Match 19.9%; Score 231.5; DB 1; Length 550;
Best Local Similarity 30.2%; Pred. No. 1.7e-13;
Matches 48; Conservative 29; Mismatches 65; Indels 17; Gaps 5;

Qy 55 SELHSHSLTCTCFSEGTGV--PAWGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGA 112
Db 386 SDLEALQKAAVAHTQGTQNTQNVQLIMQDWKFKYFISDRKKSWEAEAFVCSQA 445
Qy 113 HLNVFNTAEQNTIVQOLNESFSYFLGLSDPQGNNNQWIDKTP--YKKNVRFHILGE 169
Db 446 HLASVTSQEQALV-QITNAVVDHWIGLIDGTGEGNRWVVDGTPFDYVQSRFRWKGQPD 504
Qy 170 -----NHSAEQCASIVFWKPTGNGWINDVICETRRNSICE 203

```

RESULT 3  
LECI\_RAT  
ID LECI\_RAT STANDARD; PRT; 301 AA.  
AC P08290;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Asialoglycoprotein receptor R2/3 (Hepatic lectin 2/3) (RHL-2) (ASGP-R)  
DE (ASGPR).  
GN ASGR2 OR ASGR-2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87257885; PubMed=3600647;  
RA McPhaul M., Berg P.;  
RT "Identification and characterization of cDNA clones encoding two  
RT homologous proteins that are part of the asialoglycoprotein  
RT receptor.";  
RL Mol. Cell. Biol. 7:1841-1847(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87250656; PubMed=3597443;  
RA Halberg D.F., Wager R.E., Farrell D.C., Hildreth J., Quesenberry M.S.,  
RA Loeb J.A., Holland E.C., Drickamer K.;  
RT "Major and minor forms of the rat liver asialoglycoprotein receptor  
RT are independent galactose-binding proteins. Primary structure and  
RT glycosylation heterogeneity of minor receptor forms.";  
RL J. Biol. Chem. 262:9828-9838(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Sprague-Dawley; TISSUE=Liver;  
RX MEDLINE=89170119; PubMed=3234178;  
RA Sanford J.P., Elliott R.W., Doyle D.;  
RT "Asialoglycoprotein receptor genes are linked on chromosome 11 in the  
RT mouse.";  
RL DNA 7:721-728(1988).  
RN [4]  
RP SEQUENCE OF 201-301.  
RX MEDLINE=84111554; PubMed=6319386;  
RA Drickamer K., Mamou J.F., Binns G., Leung J.O.;  
RT "Primary structure of the rat liver asialoglycoprotein receptor.  
RT Structural evidence for multiple polypeptide species.";  
RL J. Biol. Chem. 259:770-778(1984).  
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO  
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX  
CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES  
CC TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE UNITS. AFTER LIGAND  
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND  
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE  
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE  
CC SURFACE.  
CC  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC  
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL  
CC CELLS.  
CC  
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.  
CC  
CC -1- MISCELLANEOUS: TWO TYPES OF RAT HEPATIC LECTIN HAVE BEEN  
CC IDENTIFIED, RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.  
CC RHL-2 AND RHL-3 ONLY DIFFERS IN THEIR CARBOHYDRATE STRUCTURES.  
CC  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; M16347; AAA42038.1; -;  
CC EMBL; J02762; AAA41522.1; -;

DR EMBL; X07636; CAA30476.1; -;  
DR PIR; A25417; LNRT2.  
DR PIR; A26888; A26888.  
DR PIR; A28462; A28462.  
DR PIR; A31601; A31601.  
DR HSSP; P06734; 1HLI.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS00441; C\_TYPE\_LLECTIN\_2; 1.  
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;  
KW Calcium; Signal-anchor; Phosphorylation.  
FT DOMAIN 1 58  
FT TRANSMEM 59 79  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 80 301  
FT DOMAIN 169 295  
FT DISULFID 170 181  
FT DISULFID 198 293  
FT DISULFID 271 285  
FT CARBOHYD 97 97  
FT CARBOHYD 119 119  
FT CARBOHYD 165 165  
FT CONFLICT 153 153  
FT CONFLICT 202 202  
FT CONFLICT 260 260  
SQ SEQUENCE 301 AA; 34943 MW; 3C2315E642D71279 CRC64;  
Query Match 21.5%; Score 250; DB 1; Length 301;  
Best Local Similarity 29.8%; Pred. No. 1.9e-15;  
Matches 54; Conservative 33; Mismatches 56; Indels 38; Gaps 7;  
QY 52 KRSLSELHSYHS-----SLTC-----PSEGTKVPAGCCPSKWSKFSGSSCYF 92  
DB 129 KKQKDIKADHSTLLFHLKHFPLDLTLTQLAFLFSLNGTE-----CCPVNMYVEFGSCYW 183  
QY 93 ISSEKVKWSKQNCVEMGAHLVFNTEAEQNFIVQQLNESFYFLGLSDPOGNNWOWI 152  
DB 184 FSRDGLTWAEADQCYCOMETAHLVINSREQEFVVKH-RGAFHIGWIDKDG--SWKVV 240  
QY 153 DKTPYEKNVRFWHLGEPNH-----SAPQCAASIVFWKPTGCMWMDVICETRRNSICEMN 205  
DB 241 DGETYRSNFKNAFTQPDNCOGHEEGSDCAEIL-----SDGLWMDNFCQVNRWACERK 296  
QY 206 K 206  
DB 297 R 297  
RESULT 4  
MMGL\_MOUSE  
ID MMGL\_MOUSE STANDARD; PRT; 304 AA.  
AC P49300;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage  
DE galactose/N-acetylgalactosamine-specific lectin) (MMGL).  
GN MGL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeN;  
RX MEDLINE=92268032; PubMed=1587794;  
RA Sato M., Kawakami K., Osawa T., Toyoshima S.;  
RT "Molecular cloning and expression of cDNA encoding a galactose/N-  
RT acetylgalactosamine-specific lectin on mouse tumoricidal  
RT macrophages.";  
RL J. Biochem. 111:331-336(1992).



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:59:31 ; Search time 29 Seconds  
(without alignments)  
298.916 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	22.2	301	1	LECI_MOUSE
2	252.5	21.7	311	1	LECI_HUMAN
3	250	21.5	301	1	LECI_RAT
4	244	20.9	304	1	MMGL_MOUSE
5	231.5	19.9	550	1	KUCR_RAT
6	231	19.8	207	1	LECH_CHICK
7	226.5	19.4	283	1	LECH_RAT
8	222	19.1	306	1	MMGL_RAT
9	217	18.6	283	1	LECH_MOUSE
10	200.5	17.2	290	1	LECH_HUMAN
11	194.5	16.7	147	1	LECE_ANTCR
12	189.5	16.3	548	1	KUCR_MOUSE
13	183	15.7	883	1	PGCB_MOUSE
14	183	15.7	883	1	PGCB_RAT
15	182.5	15.7	912	1	PGCB_BOVIN
16	182	15.6	1456	1	MANR_HUMAN
17	180.5	15.5	331	1	FCE2_MOUSE
18	177.5	15.2	163	1	ANP_HEMAM
19	173.5	14.9	166	1	ANP_HUMAN
20	165	14.2	146	1	TXB_TRIFL
21	163.5	14.0	173	1	LIT2_MOUSE
22	160.5	13.8	135	1	LECG_CROAT
23	158.5	13.6	165	1	LIT1_MOUSE
24	158.5	13.6	3562	1	PGCV_CHICK
25	157.5	13.5	158	1	LECG_TRIST
26	157.5	13.5	165	1	LITH_RAT
27	154.5	13.3	166	1	LITB_HUMAN
28	153.5	13.2	3381	1	PGCV_BOVIN
29	153	13.1	175	1	ANP_OSMO
30	152	13.0	123	1	ABA4_TRIAB
31	151.5	13.0	2109	1	PGCA_CHICK
32	151.5	13.0	2738	1	PGCV_RAT
33	151.5	13.0	3358	1	PGCV_MOUSE

34	151.5	13.0	3396	1	PGCV_HUMAN
35	146	12.5	125	1	BOTB_BOTJA
36	144.5	12.4	1257	1	PGCN_RAT
37	144.5	12.4	1268	1	PGCN_MOUSE
38	143	12.3	223	1	NK12_MOUSE
39	142.5	12.2	175	1	PAP1_HUMAN
40	142	12.2	263	1	KLR4_MOUSE
41	140.5	12.1	1463	1	PA2R_BOVIN
42	140.5	12.1	2124	1	PGCA_RAT
43	140.5	12.1	2132	1	PGCA_MOUSE
44	140	12.0	321	1	FCE2_HUMAN
45	139.5	12.0	549	1	LEM2_RAT

ALIGNMENTS

```
RESULT 1
LECI_MOUSE
ID      LECI_MOUSE      STANDARD;      PRT;      301 AA.
AC      P24721;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Asialoglycoprotein receptor 2 (Hepatic lectin 2) (MHL-2) (ASGP-R)
DE      (ASGPR).
GN      ASGR2 OR ASGR-2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6; TISSUE=Liver;
RX      MEDLINE=91027942; PubMed=2223888;
RA      Sanford J.P., Doyle D.;
RT      "Mouse asialoglycoprotein receptor cDNA sequence: conservation of
RT      receptor genes during mammalian evolution.";
RL      Biochim. Biophys. Acta 1087:259-261(1990).
CC      -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC      WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC      CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC      TERMINAL GALACTOSE AND N-ACETYLGLACTOSAMINE UNITS. AFTER LIGAND
CC      BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC      TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC      DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC      SURFACE.
CC      -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC      -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC      CELLS.
CC      -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC      -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X53042; CAA37211.1; -
PIR; S13165; S13165.
HSSP; P06734; IHLI.
MCD; MGI:88082; Asgr2
InterPro: IPR001304; Lectin_C.
PFam; PF00059; lectin_c; 1.
SMART; SM00034; CLECT; 1.
PROSITE; PS00615; C-TYPE LECTIN_1; 1.
PROSITE; PS50041; C-TYPE LECTIN_2; 1.
Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
Calcium; Signal-anchor; Phosphorylation.
DOMAIN 1 58
CYTOPLASMIC (POTENTIAL).
```







A;Residues: 87-98 <YU2>  
C;Comment: The functioning ligand-binding unit of this receptor is thought to be at least 100 residues long.  
C;Genetics:  
A;Gene: GDB:ASGR2; L-H2  
A;Cross-references: GDB:118755; OMIM:108361  
A;Map position: 17p13-17p11  
C;Keywords: hepatic lectin; C-type lectin homology  
C;Superfamily: hepatocellular carcinoma-associated protein  
F;1-311/Product: asialoglycoprotein receptor H2a #status predicted <MAT1>  
F;1-81-87-311/Product: asialoglycoprotein receptor H2b #status predicted <MAT2>  
F;1-58/Domin: intracellular #status predicted <INT>  
F;1-23-43-81-87-311/Product: asialoglycoprotein receptor H2c #status predicted <MAT3>  
F;59-78/Domin: transmembrane #status predicted <TM>  
F;79-311/Domin: extracellular #status predicted <EXT>  
F;177-300/Domin: C-type lectin homology <LCH>  
F;102,170,305/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 252.5; DB 1; Length 311;  
Best Local Similarity 27.0%; Pred. No. 6.5e-15;  
Matches 73; Conservative 36; Mismatches 72; Indels 89; Gaps 11;

QY 16 LSLRLWSVAGISALLS---ACFIVSCVVTYHFTYGTGKRL-SELHSYHSSLTCESE-- 59  
DB 49 LAQRLCSWVCSLLASFNLLLVIC-VTGSQSEHGRGAQQAELRLSKEAFSPSSST 107  
QY 70 -----GTVKVPWAG----- 77  
DB 108 LTEVQATSTHGGVGDKITSLGAKLEKQDLDKADHDALLFHLKHFPVDLRFVACOMELL 167  
QY 78 -----CCPASWKSFGSSCYFISSEKVKWSKSEONCVEMGAHLVVFNTAEQNFIVQ 129  
DB 168 HNSGQRTCCPVNWEHQSGWFSKGAWEAEKYLENAHLVWINSWEEQKFIQGH 227  
QY 130 LNESFSYFLGLSDPGNNWOWIDKTPYEKNVRFWHLGEPNH-----SAEQACASIVFW 182  
DB 228 TN-PFNTWGLTDSG--GSKWVDGTDYRHYNKNWAVTPDNNWHGHLGSGSDCEVEV--- 281  
QY 183 KPTGWCND-----VICETRRNSICEM 204  
DB 282 QPDG-RWNNDFCLQVRYWVCKRRNATGEV 310

RESULT 5  
JC7595  
scavenger receptor with C-type lectin type I - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C;Accession: JC7595  
R;Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.  
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001  
A;Title: Molecular cloning and functional characterization of a human scavenger receptor  
A;Reference number: JC7595; MUID:21092718; PMID:11162630  
A;Contents: Placenta  
A;Accession: JC7595  
A;Molecule type: mRNA  
A;Residues: 1-742 <NAK>  
A;Cross-references: DBBJ:AB038518  
C;Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase family, is involved in host defense. It forms a timer and plays a role in recognizing infective agents.  
C;Genetics:  
A;Gene: srcl-I  
A;Map position: 18p11.32  
C;Keywords: coiled coil; glycoprotein; transmembrane protein  
F;1-39/Domin: cytosolic (amino-terminus) #status predicted <CYT>  
F;16-19/Region: internalization signal YKRF  
F;40-56/Domin: transmembrane #status predicted <TM>  
F;57-112/Domin: extracellular #status predicted <EXT>  
F;113-335/Domin: coiled coil #status predicted <CC>  
F;369-384/Region: serine/threonine-rich #status predicted <SR>  
F;443-589/Domin: collagen-like #status predicted <COL>  
F;607-733/Domin: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 21.4%; Score 249; DB 2; Length 742;

Best Local Similarity 33.6%; Pred. No. 3.4e-14;  
Matches 49; Conservative 29; Mismatches 54; Indels 14; Gaps 4;

QY 68 SEGTKVPWAGCCPASWKSFGSSCYFISSEKVKWSKSEONCVEMGAHLVVFNTAEQNFIV 127  
DB 596 NEPTTAPEDNSCPPHKNFTDKCYFVSKEIFEAKLFCEDKSSHLVFIINTREEQWIK 655  
QY 128 QOLNESFSYFLGLSDPGNNWOWIDKTPYEKNVRFWHLGEPNH-----SAEQACASIV 180  
DB 656 KOMVGRESHWGLTDSERENEMKWLDTGTPDYKN---WKAGQPDNMGHGHGEGDCAGLI 712  
QY 181 FWKPTGWCNDVICETRRNSICEMNK 206  
DB 713 Y-----AGWNDFQCEDVNVNFIKEDR 734

RESULT 6  
JX0209  
lectin, galactose/N-acetylgalactosamine-specific - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C;Accession: JX0209; PX0009  
R;Sato, M.; Kawakami, K.; Osawa, T.; Toyoshima, S.  
J. Biochem. 111, 331-336, 1992  
A;Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylgalactose-specific for galactose 4-epimerase  
A;Reference number: JX0209; MUID:92268032; PMID:1587794  
A;Accession: JX0209  
A;Molecule type: mRNA  
A;Residues: 1-304 <SAT>  
A;Cross-references: GB:S36676; NID:g249360; PIDN:AAB22171.1; PID:g249361  
R;Oda, S.; Sato, M.; Toyoshima, S.; Osawa, T.  
J. Biochem. 104, 600-605, 1988  
A;Title: Purification and characterization of a lectin-like molecule specific for galactose 4-epimerase  
A;Reference number: PX0009; MUID:89197865; PMID:3241002  
A;Accession: PX0009  
A;Molecule type: protein  
A;Residues: 102-120;137-'X',139-151 <ODA>  
C;Superfamily: hepatic lectin; C-type lectin homology  
C;Keywords: glycoprotein; lectin; macrophage; transmembrane protein  
F;36-61/Domin: transmembrane #status predicted <TRA>  
F;173-296/Domin: C-type lectin homology <LCH>  
F;74,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.9%; Score 244; DB 2; Length 304;  
Best Local Similarity 33.7%; Pred. No. 3.6e-14;  
Matches 55; Conservative 27; Mismatches 59; Indels 22; Gaps 6;

QY 54 LSELHSYHSSLT-----FSEGTKVPWAGCCPASWKSFGSSCYFISSEKVKWSKSEONCV 108  
DB 146 VQQLRKDLKALTCQIANLKNNGSEV---ACPLHWTHEGSCYWFSEKSWFEADKYCR 202  
QY 109 EMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDPGNNWOWIDKTPYEKNVRFWHLG 168  
DB 203 LNSHLVWVNSLEQNFQNLRLANVSVW-TGLTDQ--NGPWRVVDGTFDEKFKKNWAPLQ 259  
QY 169 PNH-----SAEQACASIVFWKPTGWCNDVICETRRNSICEM 204  
DB 260 PDNWFHGLGGGEDCAHIT---TGGPWNDVDCQRTFRWICEM 298

RESULT 7  
A46274  
HIV gp120-binding C-type lectin - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 01-Dec-1995  
C;Accession: A46274  
R;Curtis, B.M.; Scharnowske, S.; Watson, A.J.  
Proc. Natl. Acad. Sci. U.S.A. 89, 8356-8360, 1992  
A;Title: Sequence and expression of a membrane-associated C-type lectin that exhibits  
A;Reference number: A46274; MUID:92390446; PMID:1518869  
A;Accession: A46274  
A;Status: preliminary  
A;Molecule type: nucleic acid



A;Title: Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor genes  
A;Reference number: S13165; MUID:91027942; PMID:2223888  
A;Accession: S13165  
A;Molecule type: mRNA  
A;Residues: 1-301 <SAN>  
A;Cross-references: EMBL:X53042; NID:g53104; PIDN:CAA37211.1; PID:g53105  
C;Superfamily: hepatic lectin; C-type lectin homology  
C;Keywords: glycoprotein; liver; Transmembrane protein  
F:170-293/Domain: C-type lectin homology <LCH>

Query Match 22.2% Score 259; DB 2; Length 301;  
Best Local Similarity 33.3%; Pred. No. 1.7e-15;  
Matches 52; Conservative 30; Mismatches 50; Indels 24; Gaps 6;

QY 63 SLTCTC-----FSEGTKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVF 117  
DB 154 TLFCQLAYFQSGNGTE-----CCPVNWFVFGGSCYFWSRDGLTWAEADQYQLENLAHLVVI 208  
QY 118 NTEAEQNFIVQOLNESFSYFLGLSDPOGNNWQWIDTPYEKNVRFVHGLGEPNH-----171  
DB 209 NSREEQDFVVKHRQSFHLWIGLTDKDG--SWKKVWDGTDYRSNYRNWAFQPDNNQHGHEQ 265  
QY 172 -SAEQCASIVFVKPTGCGWGNVDVICETRRNSICEMNK 206  
DB 266 GGEDCAEIL-----SDGHWNDNFCCQVNRWVCEKRR 297

RESULT 3  
LNRT2  
hepatic lectin 2 - rat  
N;Alternate names: asialoglycoprotein receptor RHL-2/3 (ASGP-R2/3)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-Mar-1988 #sequence\_revision 09-Apr-1998 #text\_change 22-Jun-1999  
C;Accession: B28462; A28462; A31601; A26888; A25417  
R;Halberd, D.F.; Weger, R.E.; Farrell, D.C.; Hildreth IV, J.; Quesenberry, M.S.; Loeb, J.  
J. Biol. Chem. 262, 9828-9838, 1987  
A;Title: Major and minor forms of the rat liver asialoglycoprotein receptor are independent  
A;Reference number: A28462; MUID:87250656; PMID:3597443  
A;Accession: B28462  
A;Molecule type: mRNA  
A;Residues: 1-301 <HAL>  
A;Cross-references: GB:J02762; NID:g205162; PIDN:AAA41522.1; PID:g205163  
A;Accession: A28462  
A;Molecule type: protein  
A;Residues: 88-96,'X',98-118,'X',120-129-158-177-182,'X',184,'X',186-189;192-290,'C',292-294  
R;Sanford, J.P.; Elliott, R.W.; Doyle, D.  
DNA 7, 721-728, 1988  
A;Title: Asialoglycoprotein receptor genes are linked on chromosome 11 in the mouse.  
A;Reference number: A31601; MUID:89170119; PMID:3234178  
A;Accession: A31601  
A;Molecule type: mRNA  
A;Residues: 1-301 <SAN>  
A;Cross-references: GB:X07636; NID:g57066; PIDN:CAA30476.1; PID:g57067  
R;McPhaul, M.; Berg, P.  
Mol. Cell. Biol. 7, 1841-1847, 1987  
A;Title: Identification and characterization of cDNA clones encoding two homologous proteins  
A;Reference number: A26888; MUID:87257885; PMID:3600647  
A;Accession: A26888  
A;Molecule type: mRNA  
A;Residues: 1-152,'A',154-201,'I',203-259,'C',261-301 <MCP>  
A;Cross-references: GB:M16347; NID:g206648; PIDN:AAA42038.1; PID:g206649  
A;Note: the authors translated the codon GCA for residue 153 as Arg and ATT for residue 170 as Ile.  
R;Drickamer, K.; Mamon, J.F.; Binns, G.; Leung, J.O.  
J. Biol. Chem. 259, 770-778, 1984  
A;Title: Primary structure of the rat liver asialoglycoprotein receptor: structural evidence  
A;Reference number: A25417; MUID:84111554; PMID:6319386  
A;Accession: A25417  
A;Molecule type: protein  
A;Residues: 201-259,'C',261-281,'ND',284-301 <DRI>  
C;Comment: Calcium is required for ligand binding.  
C;Superfamily: hepatic lectin; C-type lectin homology  
C;Keywords: endocytosis; glycoprotein; lectin; liver; receptor; transmembrane protein  
F:2-60/Domain: intracellular #status predicted <INT>

F:61-77/Domain: transmembrane #status predicted <TRN>  
F:78-301/Domain: extracellular #status predicted <EXT>  
F:170-293/Domain: C-type lectin homology <LCH>  
F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.8% Score 254; DB 1; Length 301;  
Best Local Similarity 29.8%; Pred. No. 4.6e-15;  
Matches 54; Conservative 33; Mismatches 56; Indels 38; Gaps 7;

QY 52 KRUSELHSYHS-----SITC-----FSEGTKVPANGCCPASWKSFGSSCYF 92  
DB 129 KKQKDIKADHSTLLFLHKHFPDLRLTLCOLAFPLSNGTE-----CCPVNWFVFGGSCYF 183  
QY 93 ISSEKVKWSKSEQNCVEMGAHLVVFTEAEQNFIVQOLNESFSYFLGLSDPOGNNWQWIDTPYEKNVRFVHGLGEPNH-----152  
DB 184 FSRDGLTWAEADQYQCMENAHLLVINSREDFVVKH-RGAFLHIWIGLTDKDG--SWKKV 240  
QY 153 DKPYEKNVRFVHGLGEPNH-----SAEQCASIVFVKPTGCGWGNVDVICETRRNSICEMN 205  
DB 241 DGYEYRSNFKNWAFTQPDNNQHGHEGSEDCAEIL-----SDGLWNDNFCCQVNRWVCEKRR 296  
QY 206 K 206  
DB 297 R 297

RESULT 4  
LNHU2A  
asialoglycoprotein receptor H2a - human  
N;Alternate names: hepatic lectin H2a  
C;Species: Homo sapiens (man)  
C;Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 10-Dec-1999  
C;Accession: A25179; A39100; B39100; I37995; A49466; B49466; S14525  
R;Spless, M.; Lodish, H.F.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6465-6469, 1985  
A;Title: Sequence of a second human asialoglycoprotein receptor: conservation of two  
A;Reference number: A25179; MUID:86016723; PMID:3863106  
A;Accession: A25179  
A;Molecule type: mRNA  
A;Residues: 1-311 <SPI>  
A;Cross-references: GB:M11025; NID:gi179080; PIDN:AAB59519.1; PID:gi179081  
R;Lederkremer, G.Z.; Lodish, H.F.  
J. Biol. Chem. 266, 1237-1244, 1991  
A;Title: An alternatively spliced minixon alters the subcellular fate of the human a



Db 182 MRDSSNPRONWNDVTCLNYFRICEM 207

Search completed: February 11, 2003, 23:02:14  
Job time : 11 secs

```

, PRIOR FILING DATE: 1999-09-15
, PRIOR APPLICATION NUMBER: PCT/US99/21547
, PRIOR FILING DATE: 1999-09-15
, PRIOR APPLICATION NUMBER: PCT/US99/23089
, PRIOR FILING DATE: 1999-10-05
, PRIOR APPLICATION NUMBER: PCT/US99/28214
, PRIOR FILING DATE: 1999-11-29
, PRIOR APPLICATION NUMBER: PCT/US99/28313
, PRIOR FILING DATE: 1999-11-30
, PRIOR APPLICATION NUMBER: PCT/US99/28564
, PRIOR FILING DATE: 1999-12-02
, PRIOR APPLICATION NUMBER: PCT/US99/28565
, PRIOR FILING DATE: 1999-12-02
, PRIOR APPLICATION NUMBER: PCT/US99/30095
, PRIOR FILING DATE: 1999-12-16
, PRIOR APPLICATION NUMBER: PCT/US99/30911
, PRIOR FILING DATE: 1999-12-20
, PRIOR APPLICATION NUMBER: PCT/US99/30999
, PRIOR FILING DATE: 1999-12-20
, PRIOR APPLICATION NUMBER: PCT/US00/00219
, PRIOR FILING DATE: 2000-01-05
, NUMBER OF SEQ ID NOS: 423
, SEQ ID NO 377
, LENGTH: 219
, TYPE: PRT
, ORGANISM: Homo Sapien
, US-09-504-011-377

```

Query Match	34.2%	Score 398.5	DB 9	Length 219
Best Local Similarity	40.8%	Pred. No. 1.1e-32		
Matches 84	Conservative 34	Mismatches 81	Indels 7	Gaps 5
Qy	3	QEQQPSQTEKRGWLS--LRLWSVAGISIALLSACFIVSCVVYTHFTYGETCKRLSELHSY	60	
Db				
Qy	5	KSFTEQCTE-RGCFSSQMFMTVACIPLFLSACFIKFCVVTFRI-FQTCDEKRFQLPEN	62	
Db				
Qy	61	HSSLTCTSEGTVPKPAWGCCPASWKSFGSSCYFISSEBKWSKSBQNCVEMGAHLVVFNTE	120	
Db				
Qy	63	PTELSCYNYGSG-SVKNCNCPILNWEYFQSSCYFFSTDTISWALSRLKNCAMGAHLVVTNSQ	121	
Db				
Qy	121	AEQFIYQQLNESYFELGLSDPQGNWNQWMDIKTPYEKNVRFWHLGEPHNSA--ECCAS	178	
Db				
Qy	122	EEQEFSLYKKPKMREFFIGLSDQVVEGQWQWGTPLTKLSLFWVDVGEPNNIATLEDCAT	181	
Db				
Qy	179	IVFKPTGQWGNVDICETRRNSICEM	204	
Db				
Qy	182	MRDSSNPRONNDVTCFLNYFRICEM	207	
Db				

```

RESULT 14
US-10-174-590-24
; Sequence 24, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Yemin
; TITLE OF INVENTION: SECRETED AND TR
; TITLE OF INVENTION: ACIDS ENCODING
; FILE REFERENCE: P343081CA2
; CURRENT APPLICATION NUMBER: US/10/1
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See Fil
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 24

```

```

; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-24

Query Match      34.2% Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.le-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps

Qy   3 QEOQPQSTKRGWLS--LRLWSVAGISIALLSACFIVSCVVYHYFTYTGETGKRLESHSY 60
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy   61 HSSUTCSEGTGVPAWGCCPASWKSFSSCFYISSEEKVMSQNCVENCAHLWVENTE 120
Db   : | | : : : : : : : | | | | | | | : : : : : | | | | | | | : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy   63 FTSLSCYNYGSG-SVKNCCPLINWEYFOSSCYFFSTDITISWALSUKNSAMGAHLWINSQ 121
Db   : | | : : : : : : : | | | | | | | : : : : : | | | | | | | : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy   121 AEQNFIYOQLNESFSYFLGLSDPQQNNWMIDKTPYEKNRVFWHLGEPNHSA--EQCAS 178
Db   : | | : : : : : : : | | | | | | | : : : : : | | | | | | | : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy   122 EEQEFLYSKPKMREFFICLSDQVEGWQWVDGTPLTLSKFWDVGEPNNIALEDCAI 181
Db   : | | : : : : : : : | | | | | | | : : : : : | | | | | | | : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~
Qy   179 IVFKWPTGWGNVDVICETRRNISCEM 204
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~
Qy   182 MRDSSNPONNDVTCLNYFRICEM 207
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~

```

	Query Match	34.2%;	Score	398.5;	DB	9;	Length	219;
	Best local Similarity	40.8%;	Pred. No.	1.1e-32;				
	Matches	84;	Conservative	34;	Mismatches	81;	Indels	7;
	Gaps							
Qy	3	QQEQPQSTKRGWL\$--LRLNWAGISIALLSACFIVSCVVTYHFTYGETGKRRLSELHSY	60					
		: :						
Dd	5	KSETCTCE-RGCFSSQMFLWTAGIPLFLSACFITRCVTFRI-FQTCDKKFKLPEN	62					
		: :						
Qy	61	HSLNCFSEBGTVPWGCGCPASMKSFSGSCYSFISSEKVSKSQNCQVMDEGAHLVVVENTE	120					
		: :						
Dd	63	FTELSCYINTGSG-SVKNCCLPMWEYFOSSCYFEFDTITSWALLKNC\$AMGAHLVVINSQ	121					
		: :						
Qy	121	AQONFIIVOOLNDSFYFTGLSDPOQNKNWQIDKTPTYEKNRFVHWILGEPNHSA-EQCAS	178					
		: :						
Dd	122	EQEQFSLYKKPKMREFFIGLSQDVVEGQMWYDGTPLTKLSFSFDVGEPNNIATLEDCAT	181					
		: :						
Qy	179	IWFVKPTGWGNWDVICETRRNSICEM	204					
		: :						

Db 63 FTLSYNTGSG-SVKNCCPLNWEYFQSSCYFFSTDTTISWALSILKNCAMGAHLVINSQ 121  
QY 121 AEONFTVOOLNESFYFLGLSDPQGNNNWQIDKTPYEKNVRFWHLGEPNHS--EOCAS 178  
Db 122 EQEFLSYKKPKRMREFFGLSDQVVEGQWVDCTPLTKLSFWDVGEPPNIIATLEDCAT 181  
QY 179 IVFWKPTGWNWDVICETRNRISICEM 204  
Db 182 MRDSSNPQWNDVTCFLNYFRICEM 207

## RESULT 12

US-09-907-841-377

; Sequence 377, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,841  
; PRIOR FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-11-29  
; Remaining Prior Application data removed - See file Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 377  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-09-907-841-377

Query Match 34.2%; Score 398.5; DB 9; Length 219;  
Best Local Similarity 40.8%; Pred. No. 1.1e-32;  
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;  
QY 3 QQCQPSQSTKRGWLS--LRLSVAGISALISACFVSCVVTYHYETGKRLSELUSY 60  
Db 5 KSETQCTE-RGCFSSQMFATVAGIPILFSLACFTRCVVTFRI-FQTCDEKRFQLPEN 62  
QY 61 HSLTCTFSEGTKVPAMGCPASWKSFGSCYFISSEKVKWSKQNCVEMGAHLVVFENTE 120  
Db 63 FTLSYNTGSG-SVKNCCPLNWEYFQSSCYFFSTDTTISWALSILKNCAMGAHLVINSQ 121  
QY 121 AEONFTVOOLNESFYFLGLSDPQGNNNWQIDKTPYEKNVRFWHLGEPNHS--EOCAS 178  
Db 122 EQEFLSYKKPKRMREFFGLSDQVVEGQWVDCTPLTKLSFWDVGEPPNIIATLEDCAT 181  
QY 179 IVFWKPTGWNWDVICETRNRISICEM 204  
Db 182 MRDSSNPQWNDVTCFLNYFRICEM 207

## RESULT 13

US-09-904-011-377

; Sequence 377, Application US/09904011  
; Publication No. US20030003530A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/904,011  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/665,350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090

```
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-902-853-377

Query Match 34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.1e-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLWSVAGISIALLSACFTIVSCVVTYHFTYGTGKRLSELHSY 60
Db 5 KSSETQCTE-RGCFSSOMFLWTVAGIPILFSLACFTITRCVVTFRI-FQTCDEKFKQLPEN 62
QY 61 HSSLTFCFSEGTKVPAWGCPASWKSFGSSCYFISSEKVKWSQNCQVEMGAHLVVFNTE 120
Db 63 FTLSYNYGSG-SVKNCCPLNWEYFQSSCYFTSTDTISWALS LKNC SAMGAHLVYVINSQ 121
QY 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNQWIDKTPYKVRWHLGEPNLSA--PQCAS 178
Db 122 EEQEFLLYKPKMREFFIGLSDQVGVQWQWVDGTPLT*LSLFDWVGEPNNIATLEDCA 181
QY 179 IVFWKPTGWGNDVICETRNESCIM 204
Db 182 MRDSSNPRQNWNDVTCFLNFRICEM 207

RESULT 11
US-09-907-824-377
; Publication 377, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherran
; APPLICANT: Gao, Wei-Qiang
```

; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-377

Query Match 34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.le-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHFTYGETCKRLSELHSY 60
DB 5 KSEETQCTE-RGCFSSQMFMTVAGIPILFLSACFITRCVVTFRF-FQTCDEKKFQLPEN 62

QY 61 HSSLTFCSEGTKVPAMGCCPASHKSFSSCYFISSEKVKMSKSEONCVEMGAHLVVFNTE 120
DB 63 FTELSCYNYGSG-SVKNCCPLNWEYFQSSCYFFSTDTISWLSLKNCSSAMGAHLVWINSQ 121

QY 121 AEONFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHS--EQCAS 178
DB 122 EQEFLSYKKPKMREFFIGLSQDVVEGQWQVGDGTPLTSLSFWDVGEPPNNTATLEDCAT 181

QY 179 IVFWKPTGCGWMDVICETRRNSICEM 204
DB 182 MRDSSNPRQWMDVTCFLNYFRICEM 207

RESULT 9
US-09-965-529-3
; Sequence 3, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529

; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc\_feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1521513CD1
US-09-965-529-3

Query Match 34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.le-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHFTYGETCKRLSELHSY 60
DB 5 KSEETQCTE-RGCFSSQMFMTVAGIPILFLSACFITRCVVTFRF-FQTCDEKKFQLPEN 62

QY 61 HSSLTFCSEGTKVPAMGCCPASHKSFSSCYFISSEKVKMSKSEONCVEMGAHLVVFNTE 120
DB 63 FTELSCYNYGSG-SVKNCCPLNWEYFQSSCYFFSTDTISWLSLKNCSSAMGAHLVWINSQ 121

QY 121 AEONFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHS--EQCAS 178
DB 122 EQEFLSYKKPKMREFFIGLSQDVVEGQWQVGDGTPLTSLSFWDVGEPPNNTATLEDCAT 181

QY 179 IVFWKPTGCGWMDVICETRRNSICEM 204
DB 182 MRDSSNPRQWMDVTCFLNYFRICEM 207

RESULT 10
US-09-902-853-377
; Sequence 377, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048

